

**EXPLORING APOBEC3 DOMAIN STRUCTURE: IMPACT ON
RECOGNITION OF DNA DAMAGE**

by Erynn Button

A thesis submitted to the School of Graduate Studies

In partial fulfillment of the requirements for the degree of

Master of Science in Medicine (Immunology and Infectious Diseases)

Faculty of Medicine

Memorial University of Newfoundland

October 2020

St. John's

Newfoundland and Labrador

Abstract

Activation-induced deoxycytidine deaminase (AID) and apolipoprotein B mRNA editing enzyme, catalytic polypeptide 3 (APOBEC3, or A3) enzymes are a family of deoxycytidine (dC) deaminases involved in somatic hypermutation (SHM), antibody class switch recombination, and antiviral responses. These enzymes share similar structural features in their single active deoxycytidine deaminase (CD) domain, but some family members (A3B/DE/F/G) possess an additional second regulatory CD domain also known as an N-terminal domain (Ntd). AID/A3s mutate the genome indiscriminately, and A3A, A3B, and A3H haplotype II have been associated with cancer development. Thus, to date the paradigm has been that AID/A3s are pro-tumour factors because they transform healthy DNA into damaged and mutated DNA. However, I observed that the activity of A3A (single domain) and A3B (double domain) is increased if their substrate DNA contains environmentally damaged nucleotides such as 8-Oxo-2'-deoxyguanosine (8oxoG) directly adjacent to the dC (in the -2 or -1 nucleotide position). When exposed to substrates containing damaged nucleotides in positions further downstream, further upstream, or containing a greater concentration of damaged bases, A3A failed to demonstrate an activity increase relative to undamaged and -1 position 8oxoG control substrates whereas A3B exhibited significantly increased activity for all distally damaged substrates compared to controls. I propose that the regulatory Ntd of A3B is a key structural component that accounts for binding distally damaged DNA substrates. The notion that AID/A3s can target pre-damaged DNA in addition to non-damaged healthy DNA is a novel insight in the field; understanding the basis of this phenomenon from both the substrate sequence and enzyme structure point of view would advance our basic knowledge and permit exploration of targeting AID/A3 enzymes in cancer.

Acknowledgements and Conflict of Interest Declaration

I wish to express my utmost gratitude to my supervisor Dr. Mani Larijani for providing me this opportunity to complete my M.Sc. This work would never have been completed if not for his advice, facilities, and funding; his belief in me since the beginning and his support during times of adversity will never be forgotten. I would also like to extend my thanks to committee members Dr. Rod Russell and Dr. Michael Grant for their support and review of my projects. Financial support was provided by Memorial University of Newfoundland and by the Natural Science and Engineering Research Council of Canada (NSERC). I declare that they have no conflicts of interest with the contents of this article.

This manuscript details my independent experiments undertaken as part of the Diamond *et al.* damaged base paper, the design of truncated and chimeric A3 enzymes to evaluate how interaction with damaged bases is affected by A3 structural features, an analysis of troubleshooting the synthesis of the truncated and chimeric A3s, and the design of a novel *in vitro* system to induce DNA base damage in cell lines using environmental mutagens. Cody Diamond (M.Sc. and M.D. candidate) and Junbum Im (Honours undergraduate student, now Ph.D. student) designed and carried out all proximally damaged base enzyme experiments and wrote and edited the manuscript. The manuscript forms the basis for my work, with minor contributions from the other authors, including myself. I (M.Sc. and M.D. candidate) assisted with the proximally damaged enzyme assays and editing of the manuscript. I conceptualized and designed the distance damage enzyme experiments and the double-domain trials and carried out the distance damage enzyme assays, which are described in this thesis. David Nicolas

Huebert and Justin John King, two Ph.D. students from our lab, designed and performed the enzyme:substrate docking structural analysis. Faeze Borzooee, a PhD student from our lab, provided technical assistance for the enzyme experiments, enzyme expression and purification. Hala Abdouni, a past Masters student from our lab, assisted with the design of substrates and optimization of experimental procedures and carried out preliminary experiments with substrates containing damaged bases. Lesley Berghuis, our Research Assistant, assisted with construction of A3A expression vectors. A previous summer student (Erin McCarthy) carried out preliminary experiments on damaged bases. Lesley Berghuis and Heather Fifield (Research Assistants) provided technical assistance in enzyme expression and purification. Lesley Berghuis provided further assistance in construction of the chimeric A3A/B expression vectors.

Proximally Damaged Base Manuscript That Forms the Basis for This Thesis

AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation

Cody P. Diamond, Junbum Im, Erynn A. Button, David N.G. Huebert, Justin J. King, Faeze Borzooee, Hala S. Abdouni, Lisa Bacque, Erin McCarthy, Heather Fifield, Lesley M. Berghuis, Mani Larijani

Biochim Biophys Acta Gen Subj. 2019 Nov;1863(11):129415.
doi: 10.1016/j.bbagen.2019.129415. Epub 2019 Aug 9. PMID: 31404619

Table of Contents

Abstract	i
Acknowledgements and Conflict of Interest Declaration	ii
Table of Contents	iv
List of Tables	vi
List of Figures	vi
Abbreviations	viii
Chapter 1 - Introduction	1
<i>Introduction to AID/APOBECs</i>	1
<i>AID/APOBEC Activity on Pre-Damaged DNA</i>	3
<i>Structural Features of AID/A3s</i>	13
<i>Rationale and Hypotheses</i>	17
<i>Experimental Objectives</i>	20
<i>Significance and Implications</i>	21
Chapter 2 - Materials and Methods	22
<i>Enzyme Expression and Purification</i>	22
<i>Preparation of Substrates and the Alkaline Cleavage Enzyme Assay</i>	29
<i>Design of Ntd-A3B, Ctd-A3B, and BNAC (Ntd-A3B, A3A as the “Ctd”) Chimera Sequences</i>	33
<i>PCR and Cloning of Hybrid Chimeric A3A/B and Truncated A3B Enzymes</i>	35
<i>Methods for Data Collection and Analysis</i>	38
Chapter 3 - Results	39
<i>Substrates Containing Distal 8oxoGs Significantly Affect A3B but Not A3A Activity, and Minimally Affect AID Activity</i>	39
<i>Design of Double Domain Damage Experiments</i>	46
<i>Generation of Hybrid Chimeric A3A/B and Truncated A3B Enzymes</i>	49
<i>Predicted Hypothetical Results of A3A/B Domain Chimeras</i>	52
Chapter 4 - Discussion	54
<i>Discussion</i>	54
Chapter 5 - Proposal for Future Work	63
<i>Summary of Proposed Future Work</i>	63
<i>Phase I – Repetition of the Distance Damage Base Experiments and Exploring A3A/B Activity When Exposed to a Range of 8oxoG Locations Within the Substrate</i>	64
<i>Phase II – Double Domain Damage Experiments with A3A and A3B</i>	65
<i>Phase III – Determination of Truncated and Chimeric A3A/A3B Activity on a Variety of Proximally Damaged Substrates</i>	66
<i>Phase IV – Synthesis and Analysis of Truncated A3F/A3G Activity on Proximally and Distally Damaged Substrates</i>	67
<i>Phase V – Development of More Sophisticated In Vitro and Ex Vivo Models for Induction of DNA Damage by Environmental Factors</i>	68
<i>Proposed Specific Protocol for Phase V</i>	69

Proposed Specific Protocol - <i>Induction of DNA Damage in Cells with Environmental Mutagens</i>	69
Proposed Specific Protocol - <i>Design of a Uridine Assay: Isolation and Preparation of Cellular DNA</i>	70
Proposed Specific Protocol - <i>Design of a Uridine Assay: Membrane Hybridization</i>	70
Proposed Specific Protocol - <i>Design of a Uridine Assay: Preparation of Uracil Standards</i>	72
Bibliography	73

List of Tables

Table 1: Catalytic parameters of AID, APOBEC3A and APOBEC3B on damaged base motifs compared to normal favored sequence motifs

Table 2: Catalytic parameters of AID, APOBEC3A, and APOBEC3B on distally damaged 8oxoG-containing motifs compared to normal favoured sequence motifs and proximally damaged 8oxoG sequence motifs

List of Figures

Figure 1: Selection of Damaged Bases and Design of Experimental Substrates for Damaged Base Trials

Figure 2: The Effect of Various Damaged DNA Bases on AID

Figure 3: The Effect of Various Damaged DNA Bases on A3A

Figure 4: The Effect of Various Damaged DNA Bases on A3B

Figure 5: Three-Dimensional Representation of AID/A3 Structural Features

Figure 6: A Linear Representation of AID/A3 Domains

Figure 7: Surface regions of AID, A3A and A3B are predicted to interact robustly with damaged bases

Figure 8: Design of Ntd-A3B, Ctd-A3B, and BNAC

Figure 9: Sequences of Ntd-A3B, Ctd-A3B, and the BNAC chimera designed for insertion into GST-containing pcDNA 3.1/V5-His TOPO vector

Figure 10: SDS Gel to Verify Purity of GST-A3A and GST-A3B Utilized for Experimentation

Figure 11: Workflow for Distance Damage Experiments

Figure 12: Model Explaining the Procedure of Alkaline Cleavage

Figure 13: Alkaline Cleavage Gels of AID Acting on ssDNA Substrates Containing 8oxoG in Various Locations and Quantities

Figure 14: Alkaline Cleavage Gels of A3A Acting on ssDNA Substrates Containing 8oxoG in Various Locations and Quantities

Figure 15: Alkaline Cleavage Gels of A3B Acting on ssDNA Substrates Containing 8oxoG in Various Locations and Quantities

Figure 16: The Effect of 8oxoG Position and Quantity on AID, A3A, and A3B Activity

Figure 17: Predicted results of A3A/B double domain activity on undamaged ssDNA substrates based on previous APOBEC3 chimera studies

Figure 18: Agarose gel verifying the size of Ntd-A3B and Ctd-A3B gene fragments post-initial PCR

Abbreviations

Abbreviation	Full name
A	Adenine
Ab	Antibody
AID	Activation induced enzyme
AIDS	Acquired immunodeficiency virus
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide
ARP	Aldehyde-reactive probe
A3	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide 3
BNAC	A3B N-terminal A3B C-terminal chimera
BSA	Bovine serum albumin
C	Cytidine
CD	Deoxycytidine deaminase domain
CD1	Deoxycytidine deaminase domain 1
CD2	Deoxycytidine deaminase domain 2
CSR	Class switch recombination
Ctd	Carboxylic acid-terminal domain
dC	Deoxycytidine
dNTP	Deoxynucleoside triphosphate
DSB	Double stranded breaks

DTT	Dithiothreitol
dU	Deoxyuridine
FCS	Fetal calf serum
FPLC	Fast protein liquid chromatography
G	Guanine
GST	Glutathione S-transferase
HIV	Human immunodeficiency virus
HMW	High molecular weight
HGNC	HMGO Gene Nomenclature Committee
Ig	Immunoglobulin
IPTG	Isopropyl β - d-1-thiogalactopyranoside
Kcat	Turnover number
Km	Michaelis Menten constant, [substrate] at $\frac{1}{2}$ Vmax
Kcat/Km	Enzyme efficiency
LB	Lysogeny broth
LMW	Low molecular weight
NCBI	National Center for Biotechnology Information
NNK	4-(Methylnitrosamino)-1-(3-pyridyl)- 1-butanone
Ntd	Amine-terminal domain

O4MeT	<i>O</i> 4-Methylthymosine
O6MeG	<i>O</i> 6-Methylguanosine
ORF	Open reading frames
PAH	Polycyclic aromatic hydrocarbons
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
RL	Recognition loops
SDS-PAGE	Sodium dodecyl sulphate– polyacrylamide gel electrophoresis
SHM	Somatic hypermutation
T	Thymine
TAE	Tris-acetate- Ethylenediaminetetraacetic acid
TBE	Tris-borate- Ethylenediaminetetraacetic acid
TE	Tris-Ethylenediaminetetraacetic acid
U	Uridine
UDG	Uracil-DNA glycosylase
V_{max}	Maximum rate of reaction
1MeA	1-Methyladenine
8oxoA	8-Oxo-2'-deoxyadenosine

8oxoG

8-Oxo-2'-deoxyguanosine

UV

Ultraviolet

Chapter 1 - Introduction

Introduction to AID/APOBECs

In a world filled with a variety of pathogens, the human immune system is tasked with keeping one step ahead. The immunoglobulins (Ig), also known as antibodies, produced by plasma B cells are one crucial tool by which viral and bacterial threats can be kept at bay ⁽¹⁾. Ig genes are initially assembled by random V(D)J recombination of the Ig locus in naïve B cells. While this random system creates a broad antibody repertoire, these naïve antibodies have a low affinity for specific antigens ⁽²⁾. With viruses and bacteria constantly mutating, a greater variety of Ig with higher affinity for specific antigens are needed. This need for variety often necessitates changes at the genomic level, facilitated by the polynucleotide editing enzyme activation-induced deoxycytidine deaminase (AID) which is a member of the apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC) deoxycytidine deaminases ⁽³⁾. There are eleven members in the AID/APOBEC family of enzymes: AID, APOBEC1, APOBEC2, APOBEC3, and APOBEC4; the APOBEC3 (A3) sub-group encompasses APOBEC3A (A3A), APOBEC3B (A3B), APOBEC3C (A3C), APOBEC3DE (A3DE), APOBEC3F (A3F), APOBEC3G (A3G), and APOBEC3H (A3H) haplotypes I and II ^(4,5).

AID mediates somatic hypermutation (SHM) and class switch recombination (CSR) of Ig genes, thereby further diversifying antibody repertoire ⁽⁶⁾. SHM is the mutation of exons at a rate higher than the background genome mutation rate ⁽⁷⁾, and CSR utilizes double stranded breaks (DSB) to switch regions of Ig loci upstream of exons ⁽⁸⁾. Class switch refers to the transformation of IgM and IgD into other Ig sub-types such as IgA,

IgE, and IgG, with each sub-type having a specific effector function in humoral immunity^(1,9). APOBEC1 edits the mRNA encoding apolipoprotein B (ApoB) to introduce an early stop codon and therefore produce a truncated ApoB for lipid transport⁽¹⁰⁾. The functions of APOBEC2 and APOBEC4 are currently being explored, but they have thus far not been demonstrated to possess deaminase or mutagenic activity in yeast assays, bacteria assays, or human cell assays⁽¹¹⁻¹³⁾. The APOBEC3 (A3) branch members (A3A-H) mutate the genomes of invading retroviruses and thus function as anti-viral agents⁽¹⁴⁾. AID/APOBECs can all deaminate any given deoxycytidine (dC) in a given DNA sequence to deoxyuridine (dU); however, each enzyme favours mutating dC in the context of signature trinucleotide “hotspots”. AID preferentially targets WRC (W=A/T and R=A/G), A3G CCC, and other A3s TTC^(15,16).

Pathologically, AID, A3A, and A3B have been shown to promiscuously mutate genome-wide. In addition to genome mutations, their activities also lead to tumour-driving chromosomal translocations which result from the double-strand DNA breaks generated by the DNA repair system attempting to repair closely spaced AID/A3-mutated dCs on opposite strands⁽¹⁷⁻²¹⁾. The result is a variety of cancers⁽²²⁾. AID has been demonstrated to play a role in leukemias, lymphomas, plasmacytomas⁽²³⁾, lung cancers⁽²⁴⁾, and gastric cancers⁽²⁵⁾. A3A has been associated with breast cancer⁽¹⁸⁾. A3B footprints have been observed in breast,⁽¹⁷⁾ bladder, and lung cancers^(18,26,27).

AID/APOBEC Activity on Pre-Damaged DNA

Normal DNA is composed of four bases: Guanine (G), Adenine (A), Cytidine (C), and Thymine (T). These are further classified as purines (possessing two hydrocarbon rings) and pyrimidines (possessing one hydrocarbon ring) (Figure 1). For the purposes of this work, damaged DNA refers to nitrogenous DNA bases that have been modified via oxidation or alkylation. Certain bases, the purines G and A, are especially susceptible to oxidative damage due to their active 8th prime carbon having a low redox potential. As a result, 8-Oxo-2'-deoxyguanosine (8oxoG) and 8-Oxo-2'-deoxyadenosine (8oxoA) are the most common damaged bases present in the human genome (²⁸⁻³⁸) (Figure 1). 8-Oxo-2'-deoxyguanosine (8oxoG) in particular can be formed through exposure to natural metabolites such as reactive oxygen species (²¹). On the other hand, exposure to the nitrosamines (ex. NNK) in tobacco smoke can cause the addition of alkyl groups to the ring nitrogen or exocyclic oxygen. *O*4-Methyl-deoxythymidine (O4MeT) and *O*6-Methyl-deoxyguanosine (O6MeG) are examples of frequently occurring alkylated bases formed in this fashion (³⁹⁻⁴⁴). 1-Methyl-deoxyadenosine (1MeA) is generated when a methyl group is transferred to the N1 nitrogen atom of deoxyadenosine, and can lead to critical replication termination (⁴⁵⁻⁴⁷). Other examples of modified DNA bases exist due to exposure to certain environmental factors. Acetaldehyde, a metabolite of ethanol, has also been known to cause a variety of DNA adducts, predominantly *N*(2)-ethylidenedeoxyguanosine (^{48,49}). Other oxidation and methylation reactions can be catalyzed by ultraviolet (UV) radiation, dietary nitrosamines, and the polycyclic aromatic hydrocarbons (PAH) of cigarette smoke (⁵⁰⁻⁵²).

Many of these damaged DNA bases can be formed after exposure to chemotherapeutic drugs, as their mechanisms rely on forcing replication termination in the DNA of cancer cells (⁵³⁻⁵⁶) (Figure 1). Damaged bases have also been found in abundance in both normal genomes and tumours where AID/APOBEC expression is increased (⁵⁷⁻⁶⁰). Specifically, it is well established that 8oxoG formed by reactive oxygen species (ROS) is present at elevated levels in tumour cells (⁶¹⁻⁶⁷) (Figure 1). ROS are commonly produced upon exposure to ionizing radiation and chemotherapeutic agents, hyperthermic conditions, when antioxidant enzymes are inhibited, or when NADPH and glutathione stores are depleted (⁶¹). While they are a part of normal biological processes such as intracellular oxygen metabolism, immune-mediated attack of pathogens, signal transduction, and gene expression, they are also associated with pathological processes including ageing (⁶⁸), Alzheimer's disease (⁶⁹), diabetes mellitus (⁷⁰), Behçet's disease (⁷¹⁻⁷³), and Sjögren's syndrome (⁷⁴⁻⁷⁶).

A

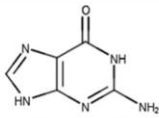
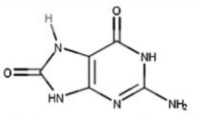
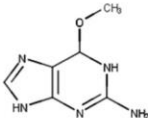
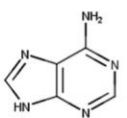
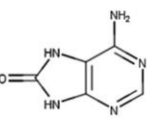
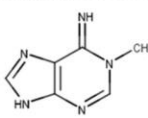
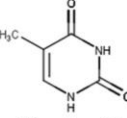
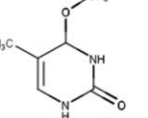
Normal Base	Damaged Base	Brief Description & Rationale of Damaged Base
 <p>Guanine (G)</p>	 <p>8-Oxoguanine (8oxoG)</p>	<p>8oxoG is the most abundant damaged nucleoside caused by the oxidation of guanosine by ROS. A common biomarker of endogenous oxidative stress, 8oxoG is responsible for G//A mismatches in DNA due to its resemblance to thymine. Clustered events of guanosine oxidation and also lead to double-stranded breaks (DSBs).</p>
	 <p>O6-Methylguanine (O6MeG)</p>	<p>O6MeG is another product of alkylation at the O6 atom of guanosine, resulting in a dual binding affinity for both cytosine and thymine during DNA replication. O6MeG is highly mutagenic and toxic due to its evasion of proofreading through steric mimicry when bound to DNA polymerase. The primary defense against O6MeG is O6-Methylguanine-DNA methyltransferase (MGMT), thus making this a promising target in cancer therapy.</p>
 <p>Adenine (A)</p>	 <p>8-Oxoadenine (8oxoA)</p>	<p>8oxoA is the most common consequence of adenine oxidation through hydroxyl radicals and DNA irradiation. 8oxoA can induce A>C mutations, although at approximately a four-fold lower frequency than those induced by 8oxoG. When occurring in tandem with an apurinic/aprimidinic (AP) site, 8oxoA is also capable of inhibiting incision of that AP site by DNA glycosylases and AP endonucleases.</p>
	 <p>1-Methyladenine (1MeA)</p>	<p>1MeA is a result of adenosine alkylation at N1. Although not highly mutagenic, 1MeA lacks Watson-Crick (W-C) base pairing capability and thus blocks DNA polymerases during DNA replication. Adenosine alkylation occurs through reactive alkylating agents such as dimethylsulfide and endogenous aldehydes.</p>
 <p>Thymine (T)</p>	 <p>O4-Methylthymine (O4MeT)</p>	<p>O4MeT is caused by alkylation damage at the O4 of thymine. Like 1MeA and O6MeG, this nucleotide is also highly mutagenic and carcinogenic, having lower frequencies of repair than the other alkylated bases. Modification of the O4 results in a shift in favorability from adenine to guanosine, causing T>C transitions.</p>



Figure 1: Selection of Damaged Bases and Design of Experimental Substrates for Damaged Base Trials

a) List of all damaged bases used in this study. From left to right, name of each base, the base chemistry relative to the normal version, and a brief description and rationale for selection. The brief statements on the attributes of each base are summarized from cited references (^{31,32,41,53,54,56,77-97}).

b) Standard bubble and single-stranded substrates used in the alkaline cleavage deamination assay. The upper panel shows an example of the partially single-stranded (bubble) substrates used for testing the enzymatic activity of AID, and the lower panel shows fully single stranded substrates used for testing the activities of A3A and A3B. The catalytic efficiency of each enzyme on substrates containing the damaged base (denoted by “X”) in the -1 or -2 position (Right Panels) was compared against a control substrate that is identical except for containing the normal undamaged version of the same base in the same -1 or -2 position (Left Panels).

***Adapted by permission from Elsevier: Elsevier. Biochimica et Biophysica Acta (BBA) - General Subjects. AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². © 2019 Elsevier B.V. All rights reserved. (2019)*

Since tumour cell genomes have an abundance of these damaged bases, and since AID/APOBEC expression is also often upregulated in tumours, our lab sought to examine whether and to what extent AID/APOBECs can recognize or act on DNA sequences that contain these damaged bases. Since AID/APOBECs have a demonstrated specificity for the -2 and -1 position bases relative to their target dC, substrates were generated in which the -2/-1 bases were composed either of normal DNA sequences, which were the most favoured AID/A3 sequences described to date, or in which the -2/-1 bases were replaced by the aforementioned damaged bases. Our research team found that AID, A3A, and A3B deaminate dCs located within damaged trinucleotide sequences (ex. T(8oxoA)C) with equivalent or increased efficiency compared to dCs within normal DNA (ex. TAC) (Figures 2-4).

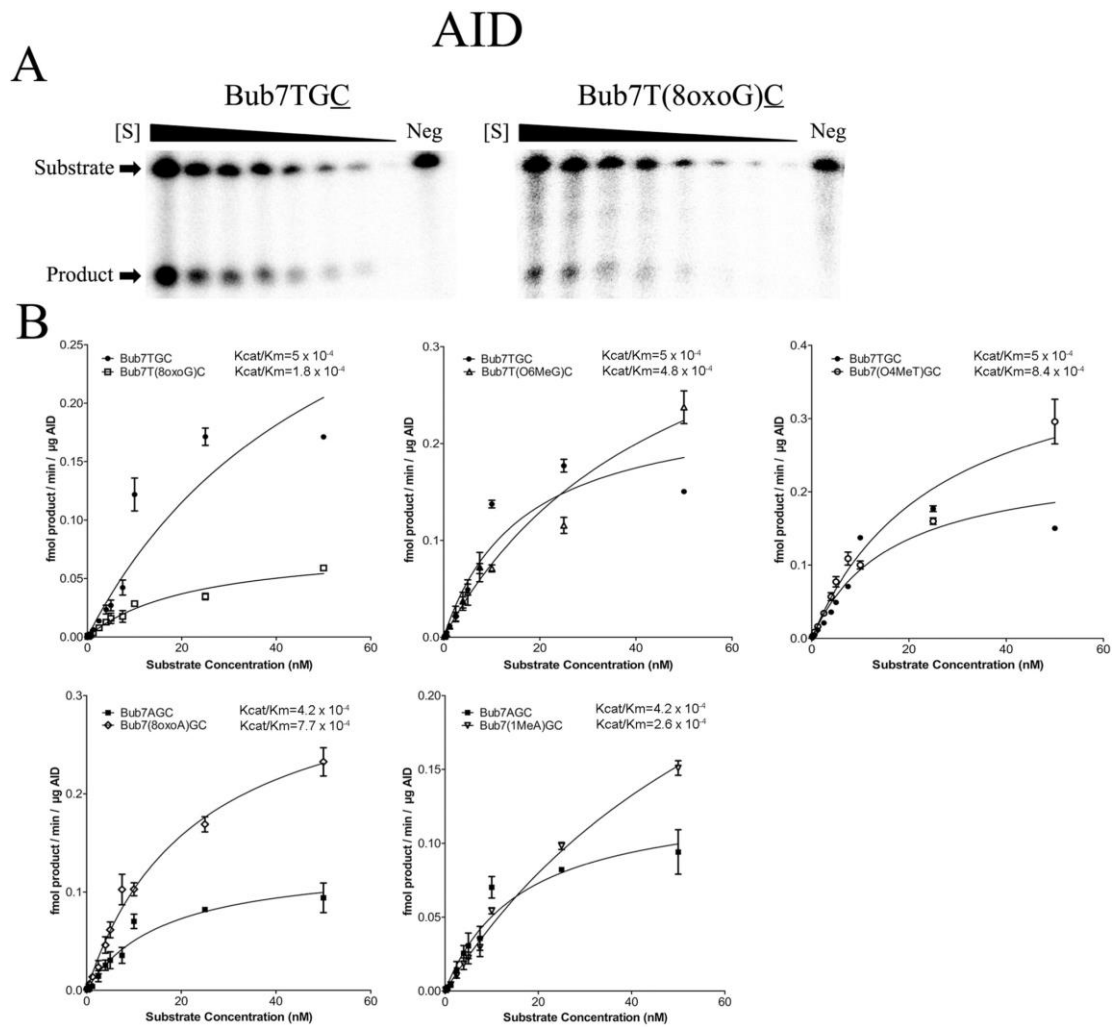


Figure 2: The Effect of Various Damaged DNA Bases on AID

a) Alkaline cleavage gels of AID kinetics on undamaged single-stranded TGC and damaged single-stranded T(8oxoG)C substrates. Substrate concentrations from left to right are 75, 50, 40, 25, 12.5, 6.25, 3.125, and 1.56 fmol, followed by a no-enzyme negative control (Neg).

b) AID was shown to be active on bubble DNA substrates containing damaged DNA bases. AID demonstrated increased catalytic activity on substrates containing O4-methylthymosine (O4MeT), O4-methylguanosine (O4MeG), 8-oxoadenosine (8oxoA), and 1-methyladenosine (1MeA), relative to their undamaged controls. Of these, 8oxoA displayed the most significant increase in activity compared to its undamaged control (TGC). AID was active on 8-oxoguanosine (8oxoG), but not more so than the undamaged TGC substrate ($n=9$).

****Reprinted from AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. 1863/11, Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². AID, APOBEC3A**

and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. Copyright 2019, with permission from Elsevier.

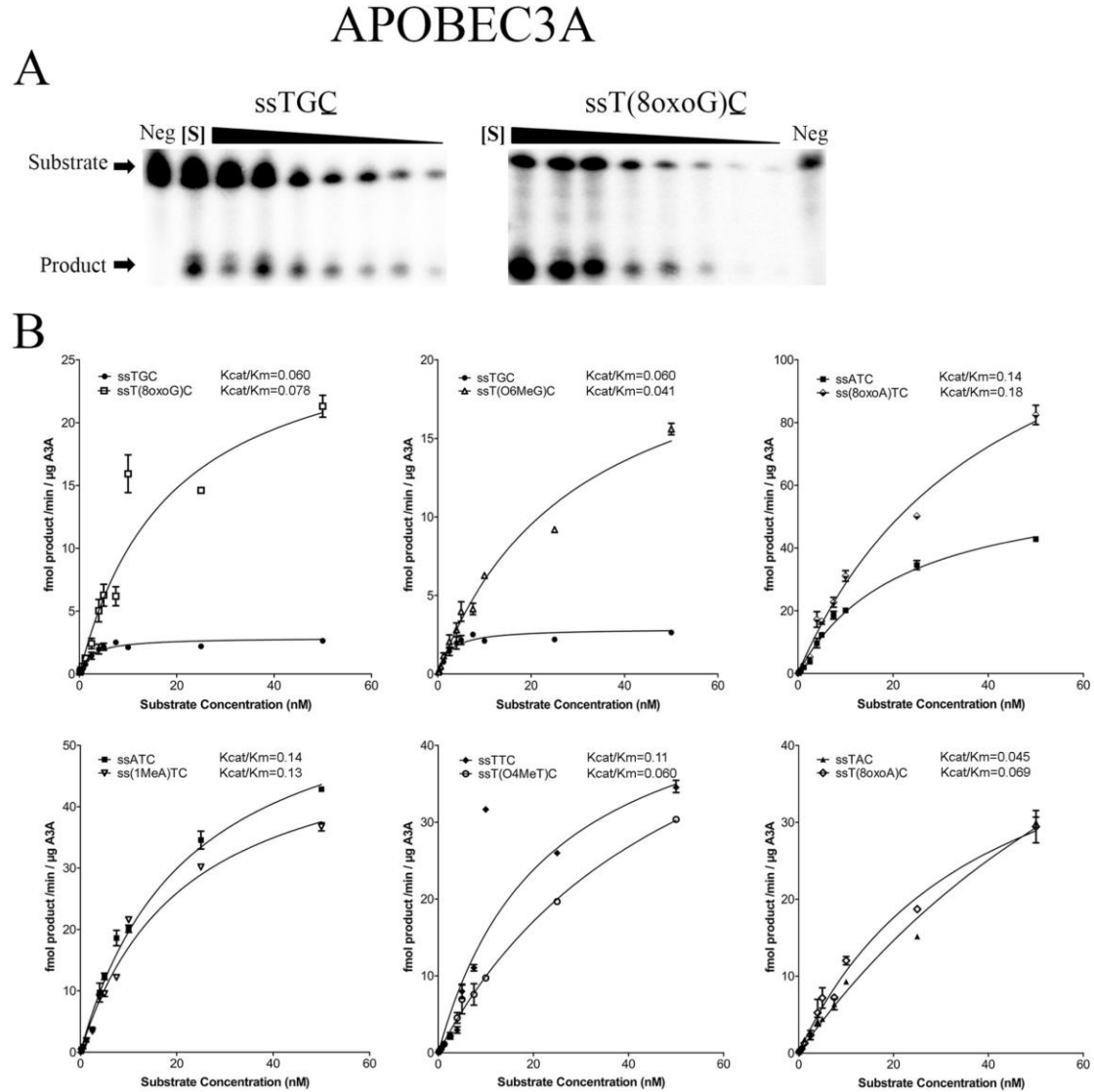


Figure 3: The Effect of Various Damaged DNA Bases on A3A

a) Alkaline cleavage gels of A3A kinetics on undamaged single-stranded TGC and damaged single-stranded T(8oxoG)C substrates. Substrate concentrations from left to right are 75, 50, 40, 25, 12.5, 6.25, 3.125, and 1.56 fmol, followed by a no-enzyme negative control (Neg).

b) A3A was shown to be active on ssDNA substrates containing damaged DNA bases. A3A demonstrated increased catalytic activity on substrates containing 8-oxoguanine (8oxoG) and 8-oxoadenosine (8oxoA), relative to their undamaged controls. Of these, 8oxoG displayed the most significant increase in activity compared to its undamaged

control (TGC). A3A was active on O4-methylguanosine (O4MeG), 1-methyladenosine (1MeA), O4-methylthymosine (O4MeT), and T(8oxoA)C substrates, but not more active than it was on undamaged controls (n=9).

****Reprinted from AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. 1863/11, Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. Copyright 2019, with permission from Elsevier.**

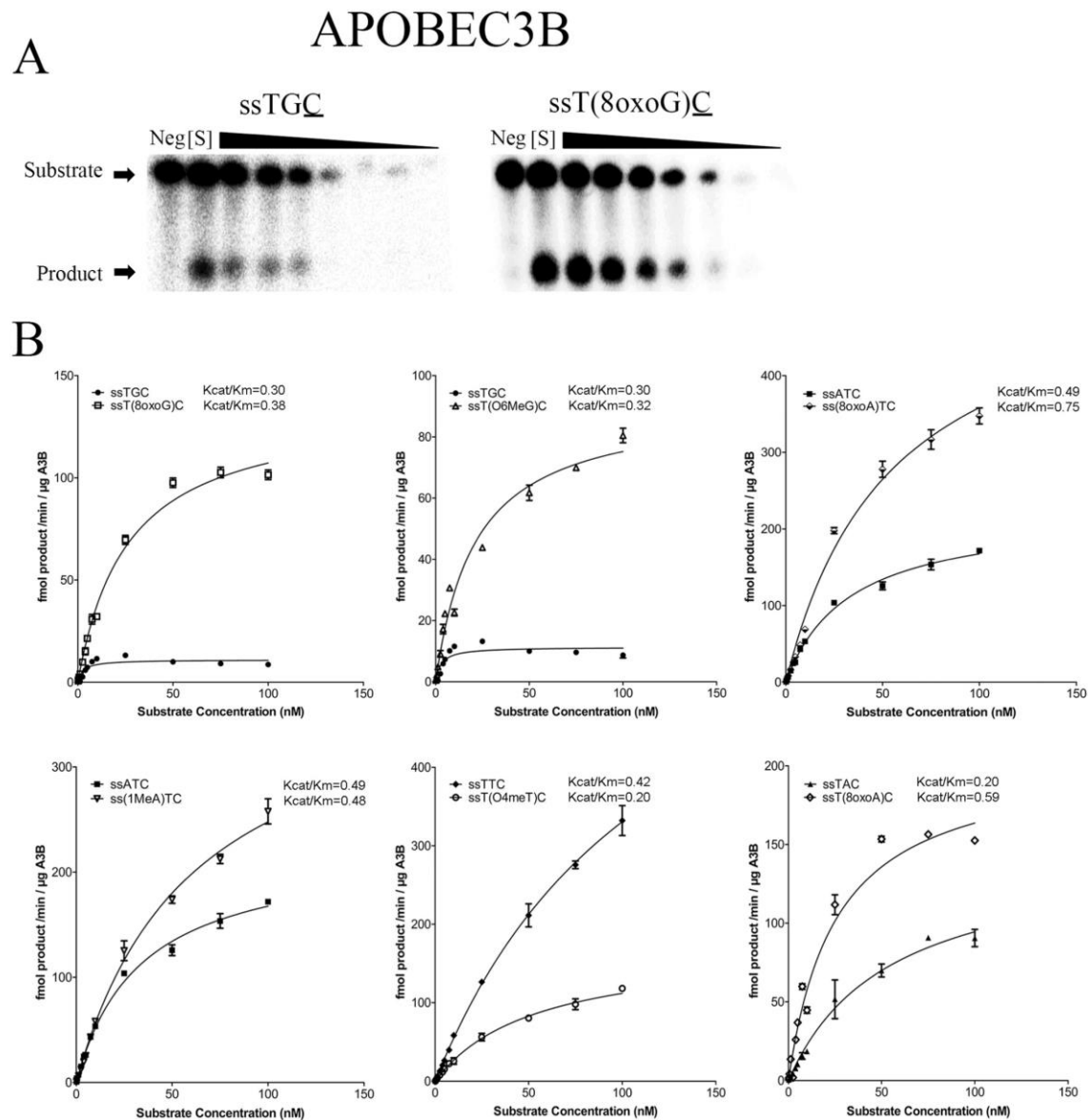


Figure 4: The Effect of Various Damaged DNA Bases on A3B

- a) Alkaline cleavage gels of A3B kinetics on undamaged single-stranded TGC and damaged single-stranded T(8oxoG)C substrates. Substrate concentrations from left to right are 75, 50, 40, 25, 12.5, 6.25, 3.125, and 1.56 fmol, followed by a no-enzyme negative control (Neg).
- b) A3B was shown to be active on ssDNA substrates containing damaged DNA bases. A3B demonstrated significantly increased catalytic activity on substrates containing 8-oxoguanine (8oxoG), O4-methylguanosine (O4MeG), 8-oxoadenosine (8oxoA), 1-methyladenosine (1MeA), O4-methylthymosine (O4MeT), and T(8oxoA)C substrates, relative to their undamaged controls. Of these, 8oxoG displayed the most significant increase in activity compared to its undamaged control (TGC) (n=9).

***Reprinted from AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. 1863/11, Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. Copyright 2019, with permission from Elsevier.*

For example, A3A was active on both TGC and T(8oxoG)C, as evidenced by Figure 3 and Table 1.

Table 1

Catalytic parameters of AID, APOBEC3A and APOBEC3B on damaged base motifs compared to normal favored sequence motifs.

Enzyme	Substrate	K_M (μ M)	K_{cat} (min^{-1})	K_{cat}/K_M ($\text{min}^{-1}\mu\text{M}^{-1}$)	K_{cat}/K_M fold difference
AID	TGC	0.035 ± 0.0071	1.7×10^{-5}	5×10^{-4}	
	T(8oxoG)C	0.023 ± 0.0023	4.1×10^{-6}	1.8×10^{-4}	0.36
	T(O6MeG)C	0.048 ± 0.0073	2.3×10^{-5}	4.8×10^{-4}	0.96
	(O4MeT)GC	0.026 ± 0.0033	2.2×10^{-5}	8.4×10^{-4}	1.7
	AGC	0.016 ± 0.0023	6.8×10^{-6}	4.2×10^{-4}	
	(8oxoA)GC	0.023 ± 0.0021	1.7×10^{-5}	7.7×10^{-4}	1.8
	(1MeA)GC	0.077 ± 0.0073	2×10^{-5}	2.6×10^{-4}	0.62
APOBEC3A	TGC	0.0024 ± 0.00051	1.5×10^{-4}	0.060	
	T(8oxoG)C	0.019 ± 0.0035	1.5×10^{-3}	0.078	1.3
	T(O6MeG)C	0.030 ± 0.0043	1.2×10^{-3}	0.041	0.68
	TTC	0.023 ± 0.0075	2.6×10^{-3}	0.11	
	T(O4MeT)C	0.053 ± 0.0062	3.2×10^{-3}	0.060	0.55
	TAC	0.10 ± 0.018	4.5×10^{-3}	0.045	
	T(8oxoA)C	0.037 ± 0.0060	2.6×10^{-3}	0.069	1.73
	ATC	0.022 ± 0.0018	3.2×10^{-3}	0.14	
	(8oxoA)TC	0.041 ± 0.0047	7.5×10^{-3}	0.18	1.29
	(1MeA)TC	0.021 ± 0.0019	2.7×10^{-3}	0.13	0.93
APOBEC3B	TGC	0.0028 ± 0.00095	8.1×10^{-4}	0.30	
	T(8oxoG)C	0.026 ± 0.002	9.9×10^{-3}	0.38	1.3
	T(O6MeG)C	0.021 ± 0.002	6.7×10^{-3}	0.32	1.1
	TTC	0.138 ± 0.030	5.9×10^{-2}	0.42	
	T(O4MeT)C	0.080 ± 0.010	1.6×10^{-2}	0.20	0.48
	TAC	0.054 ± 0.0094	1.1×10^{-2}	0.20	
	T(8oxoA)C	0.026 ± 0.0031	1.5×10^{-2}	0.59	3.0
	ATC	0.033 ± 0.0020	1.6×10^{-2}	0.49	
	(8oxoA)TC	0.055 ± 0.0054	4.1×10^{-2}	0.75	1.5
	(1MeA)TC	0.061 ± 0.0053	2.9×10^{-2}	0.48	1.0

Table 1: Catalytic parameters of AID, APOBEC3A and APOBEC3B on damaged base motifs compared to normal favored sequence motifs

***Reprinted from AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. 1863/11, Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. Copyright 2019, with permission from Elsevier.*

Moreover, when compared to undamaged controls, A3A was more active on substrates containing 8oxoG and 8oxoA (1.3 times and 1.7 times more active, respectively) (Table 1, graphs in Figure 3). Furthermore, *in silico* analysis predicted stronger binding of

damaged DNA bases (such as 8oxoG) within A3A's catalytic pocket (Figure 5), with the damaged residue being situated more directly over the catalytic zinc ion (purple). This could indicate that AID/A3s are evolutionarily well equipped to act on damaged DNA.

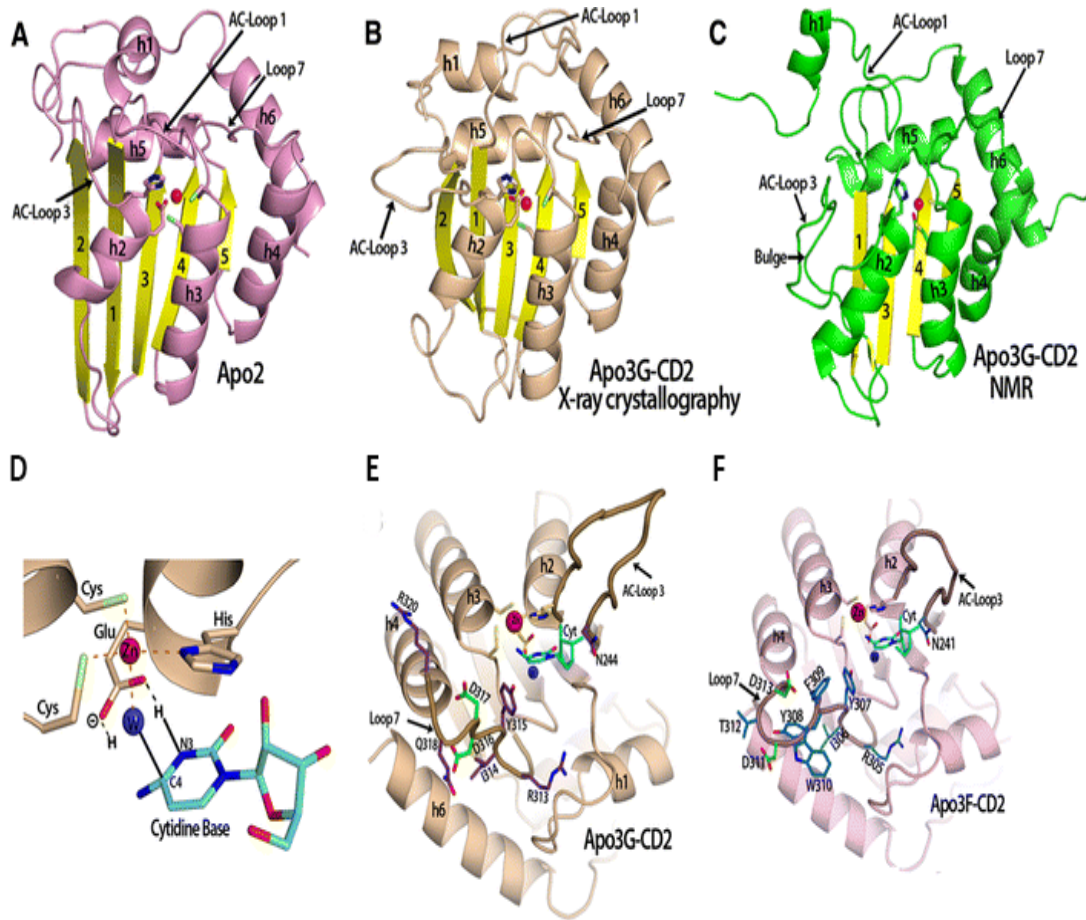


Figure 5: Three-Dimensional Representation of AID/A3 Structural Features

a,b,c) Structures representative of AID monomer and the CD2 domain of double-domain A3s, modelled here as A3G-CD2 and A3F-CD2 because of limited availability of X-ray crystallography and NMR data for AID/A3s. The five β -strands surrounded by six α -helices and seven loops are visible, with loops 1 and 3 mediating the catalytic pocket. d) Model of the catalytic pocket or active site, containing a zinc ion (Zn), two cysteine (Cys) and one histidine (His), and a catalytic glutamic acid surface (Glu), all interacting with a cytosine substrate. e) A sample binding groove for A3G-CD2.

f) A sample binding groove for A3F-CD2, demonstrating the high structural homology of APOBEC3 enzymes.

***Adapted/Translated by permission from Springer: Springer Cell and molecular life sciences – CMLS.*

The current structural and functional understanding of APOBEC deaminases.

Bransteitter R1, Prochnow C, Chen XS. © 2009 Springer. All rights reserved. (2009)

Structural Features of AID/A3s

Structurally, AID/A3s have either one or two deoxycytidine deaminase domains (CDs); AID, A3A, A3C, and A3H are single-domain (monomer), whereas A3B, A3DE, A3F, and A3G are double-domain (^{98,99}). These domains are labelled as CD1 and CD2, or the N-terminal domain (Ntd) and C-terminal domain (Ctd) (Figure 6). Single-domain A3s, such as A3A, contain one active domain called a CD, or deoxycytidine deaminase domain (^{5,100,101}). Double-domain A3s such as A3B, A3DE, A3G, and A3F contain both an active CD2 domain (the Ctd) and a regulatory CD1 domain (the Ntd) (^{5,100,101}). A3A and the CD2 (Ctd) of A3B are highly homologous (¹⁰⁰).

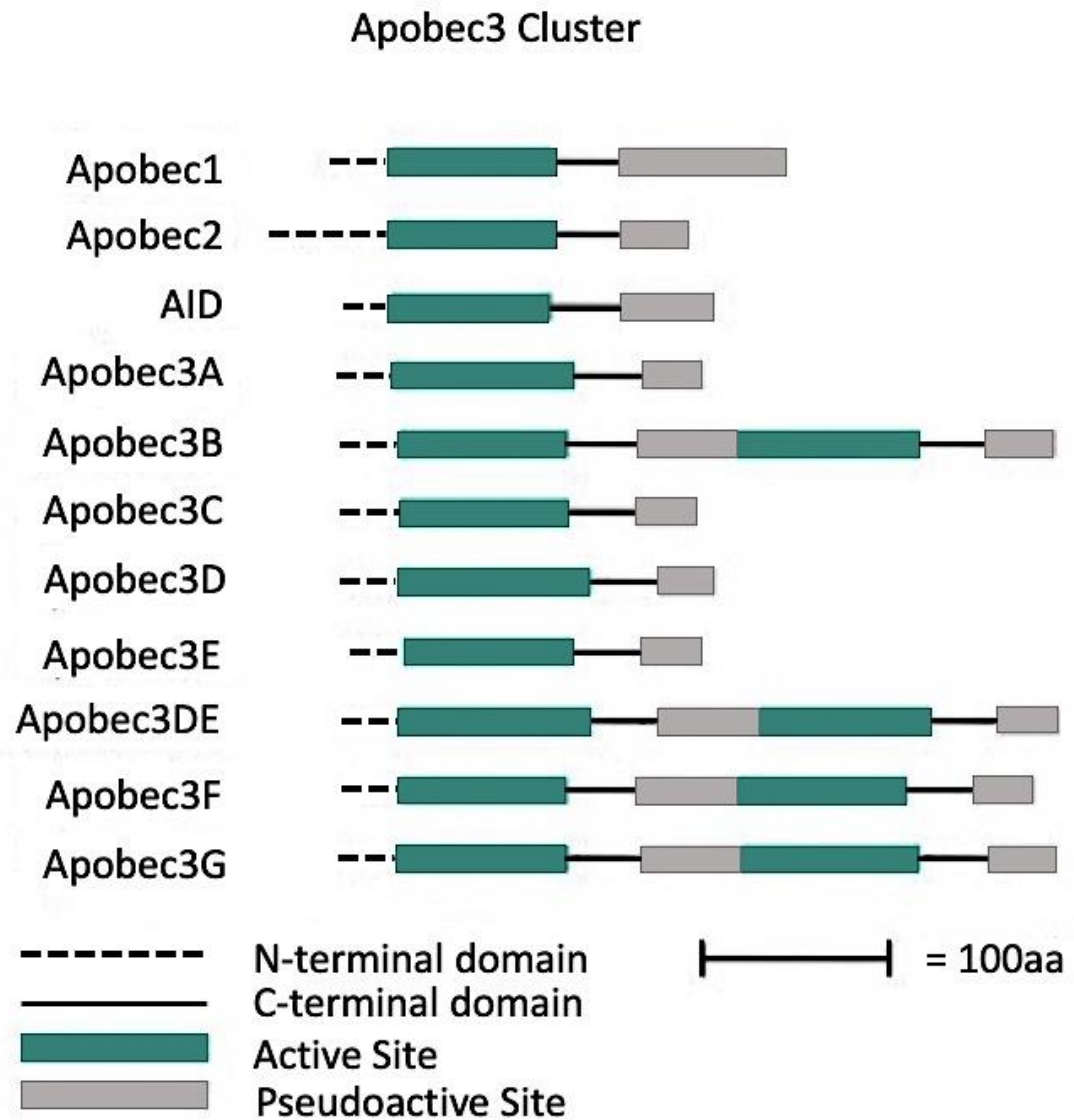


Figure 6: A Linear Representation of AID/A3 Domains

***Adapted by permission from PLoS: PLoS Biology. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. Sawyer SL, Emerman M, Malik HS. © 2004 PLoS. All rights reserved. (2004)*

The monomeric and CD2 domains are composed of five β -strands surrounded by six α -helices and seven loops (⁵). The catalytic site of these enzymes contains a zinc ion, several zinc-coordinating residues (two cysteines and a histidine for A3s or a third cysteine for AID), and a catalytic glutamic acid surface. At neutral pH AID is highly positively charged, particularly along its two ssDNA binding grooves (^{5,102}). ssDNA substrates are preferred by A3A and A3B (^{100,103}), whereas AID prefers dsDNA bubble substrates (^{104,105}). Substrate access to the catalytic (or active) site is mediated by loops 1 and 3 (⁵) (Figure 5).

The double-domain A3s, namely A3F, A3G, and more recently A3B have been observed to exist in both high molecular weight (HMW) and low molecular weight (LMW) forms in human cells (¹⁰¹). To date, there have been few studies of the biochemical and structural properties of Ntd-A3B (¹⁰¹). However, more is known about the Ntds of other double-domain A3s, namely A3G and A3F.

The Ctd of A3G is the active domain responsible for cytidine deamination. A3G possesses the best-characterized Ntd of all double-domain A3s. The known *in vitro* functions of Ntd-A3G are the enhancement of substrate binding, processivity, and deamination activity, the mediation of RNA-dependent oligomerization of A3G, and the incorporation of A3G into the HIV virion (^{14,106,107}). Before viral infection, A3G is found cytoplasmically in the HMW form. When a sufficient level of RNaseA is produced, A3G is converted to its active LMW form. Trp94 and Trp127 are the crucial residues involved in HMW complex assembly (^{14,101,108}).

In A3F, the Ctd is the active deoxycytidine deaminase domain and the Ntd is the regulatory domain. The role of the regulatory domain is to enhance catalytic activity and

mediate the assembly of HMW complexes for A3 double-domain proteins assembly^(14,101,108).

Based on limited literature, A3B appears to share structural similarities with A3G and A3F. It possesses an active deoxycytidine deaminase Ctd and a catalytically inactive Ntd that is hypothesized to be involved with nucleic acid interactions and protein oligomerization⁽¹⁰¹⁾. There is evidence that A3B is capable of forming HMW and LMW complexes similar to those of other double-domain A3s, particularly within the context of viral restriction⁽¹⁰⁷⁾. Unfortunately, Ntd-A3B is poorly soluble and interacts complexly with nucleic acids, resulting in characterization challenges⁽¹⁰¹⁾. In 2017, Xiao *et al.* identified a unique structural feature of A3B: the enzyme was active in its HMW form, and did dissociate to a LMW form under standard RNaseA treatment. They also demonstrated that W127, a residue within the Ntd, was a key residue for HMW complex formation of A3B. Furthermore, it was found that the highly positively charged surface (“positive patches” containing numerous arginine residues) of the Ntd was involved in RNA-dependent attenuation of catalysis⁽¹⁰¹⁾. A3B’s Ctd can function alone, but its activity was determined to be weaker than that of A3B containing both an Ntd and a Ctd⁽¹⁰¹⁾.

I sought to examine whether the addition of a regulatory Ntd to A3A will enhance its activity by using an A3B-A3A chimera. To date, there are limited to no published studies that added an entire domain of A3B to A3A. Several studies have replaced specific loops of A3B with loops from A3A^(100,103), and others have exchanged substrate sequence recognition loops (RL) 1 and 2 of A3A with A3G RL (and *vice versa*)⁽¹⁰⁹⁾. In the A3A/A3G chimera study, substituting RL1 of A3A for an A3G RL made a viable protein,

whereas RL2 substitution for an A3G RL did not (¹⁰⁹). A few studies have attempted to create A3 chimeras by substituting entire domains, and a common trend of enzymatic activity being determined by the Ctd can be observed amongst most of them. Goila-Gaur *et al.* replaced the Ctd of A3G with A3A, kept the A3G Ntd regulatory domain, and found that antiviral activity was still exhibited (¹¹⁰). Pak *et al.* generated A3B/A3G chimeras, but only replaced the first 60 amino acids of A3B (part of the Ntd) with an A3G sequence. This alteration retargeted A3B from the nucleus to the cytoplasm, A3G's primary location, and enhanced the enzyme's ability to restrict HIV; ergo A3B gained some characteristics of A3G and the chimera was active (¹¹¹). Haché *et al.* designed A3F/AID chimeras, finding that WRC motifs were preferred when AID was the Ctd and the TTC motifs were preferred when A3F was the Ctd. When just the Ctd of A3F was isolated, its activity was comparable to double-domain A3F and its normal TTC targets were preferred (¹⁰⁸). McDougale *et al.* created chimeras of human A3B and rhesus A3B, showing that when the Ctd was rhesus the enzyme exhibited rhesus antiviral restriction levels and when the Ctd was human the enzyme demonstrated human A3B antiviral restriction levels (¹¹²). Hakata and Landau generated human A3F/A3G chimeras, and found that whichever A3 subtype comprised the Ctd determined overall activity characteristics (¹¹³).

Rationale and Hypotheses

Our 2019 published manuscript (Diamond *et al.*), to which I contributed as a co-author, entitled "AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines

neighbouring DNA damage induced by oxidation or alkylation”, offered evidence that AID and APOBECs are active on bubble and ssDNA substrates containing damaged DNA bases. While these lesions are common in cancerous cells where repair mechanisms are often inhibited to some degree, this study offered the first insight into AID/A3 activity in these situations. When the damaged bases were located at the -1 or -2 positions relative to the target dC, AID, A3A, and A3B all exhibited several fold K_{cat} increases compared to undamaged substrates. While this was not true for every damaged substrate, some examples include AID’s 1.8-fold activity increase on (8oxoA)GC (Figure 2, Table 1), A3A’s 1.73-fold activity increase on T(8oxoA)C (Figure 3, Table 1), and A3B’s 3-fold activity increase on T(8oxoA)C (Figure 4, Table 1).

The work described above was carried out with substrates in which the damaged bases were confined to -2/-1 positions, and only with a single damaged base proximal to the target dC. *Here, my first hypothesis in regards to forthcoming unpublished data was that an increased concentration of damaged bases might further enhance AID/A3 activity, possibly by causing structural changes to the DNA substrate that could attract these enzymes.* The published work carried out thus far was restricted to wild type versions of the AID/A3. *My second hypothesis based on forthcoming unpublished data was that non-wild type versions of AID/A3s (e.g. truncated or chimeric versions) would have altered recognition of substrates with damaged DNA.* These two hypotheses deal respectively with understanding the substrate and enzyme structure aspect of this novel phenomenon of damaged DNA recognition by the AID/A3 enzymes.

Furthermore, *in silico* analysis predicted stronger binding of damaged DNA bases (such as 8oxoG) within A3A’s catalytic pocket (Figure 7), with the damaged residue

being situated more directly over the catalytic zinc ion (purple). This could indicate that AID/APOBECs are evolutionarily well equipped to act on damaged DNA.

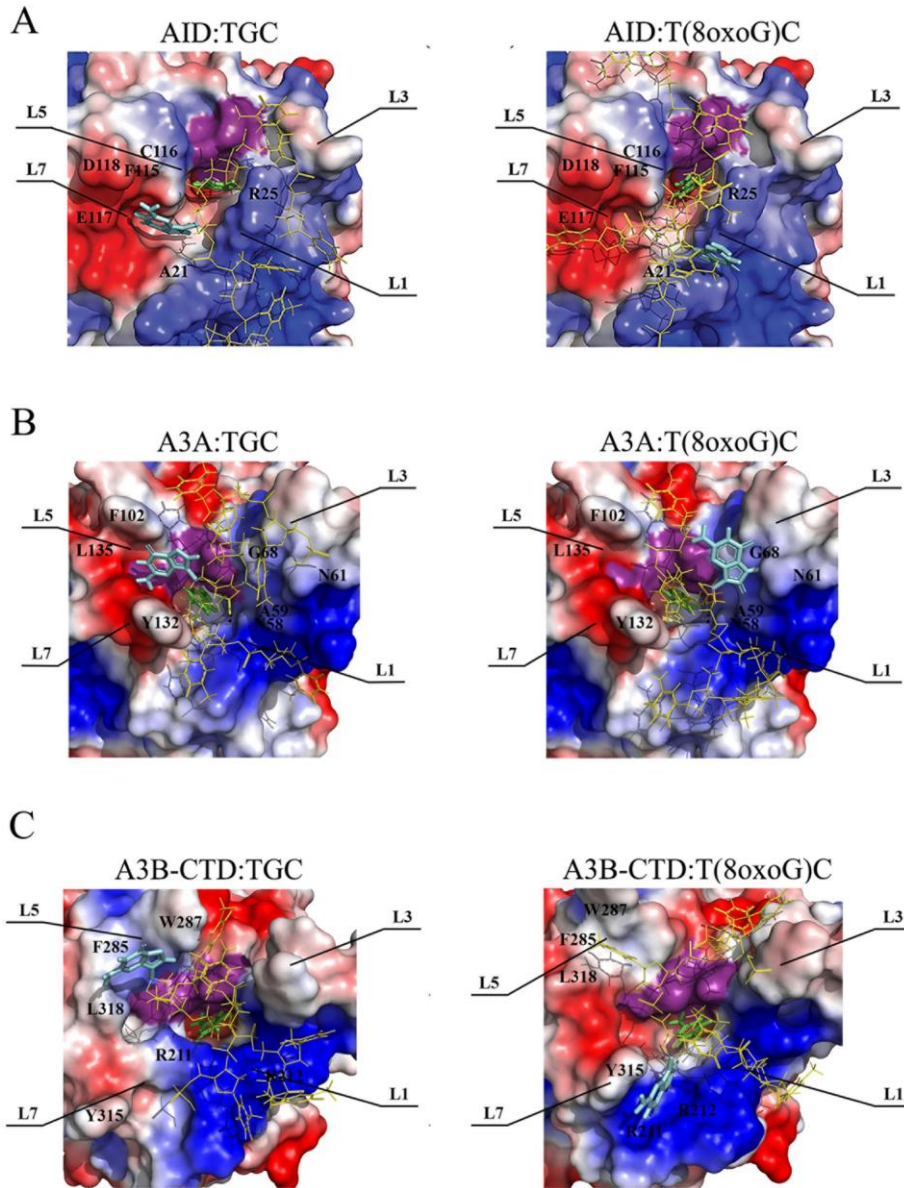


Figure 7: Surface regions of AID, A3A and A3B are predicted to interact robustly with damaged bases

AID (A), A3A (B) and A3B (C) were docked with normal and damaged substrates to examine interactions. Shown are representative analyses of interactions of the catalytic

pockets and surrounding regions of AID/A3A/A3B with TGC (Left Panels), and T(8oxoG)C (Right Panels). Residues frequently involved in nucleobase interactions are denoted. Negative and positively charged residues are colored red and blue respectively while purple represents the catalytic pocket residues. The DNA substrate is denoted by the colours yellow, green, and cyan, which respectively represent the main DNA chain, the cytidine in the catalytic pocket, and the undamaged nucleobase (-1 dG of TGC; left panels) compared to the damaged nucleobase (-1 8oxoG of T(8oxoG)C; right panels).

***Reprinted from AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. 1863/11, Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. Copyright 2019, with permission from Elsevier.*

Experimental Objectives

To test our first hypothesis that damaged bases in positions other than -2/-1 or with higher numbers would impact AID/A3 activity, substrates containing 8oxoG in positions farther from the dC ought to be designed and tested for AID/A3 activity, in comparison to control substrates with no damaged bases or with damaged bases present only in the -2/-1 positions.

To test our second hypothesis that domain structures of AID/A3s are important as regulators of their activities on damaged DNA sequences, enzymes ought to be designed wherein these domains are either truncated or exchanged to form chimeric enzymes which are then tested in comparison to their wild type counterparts for activity patterns on normal *versus* damaged DNA sequences.

Significance and Implications

The work undertaken in this thesis has both basic knowledge generation and potential future applications. A better comprehension of A3 regulatory domains might provide novel oncological drug targets to modulate or inhibit A3 activity in cancers.

As discussed previously, little to no research has addressed the impact of environmental DNA damage on A3 activity. Better understanding the extent and mechanisms of this novel aspect of A3 activity is also an important advance in our basic knowledge of how these enzymes function at the biochemical and genomic levels.

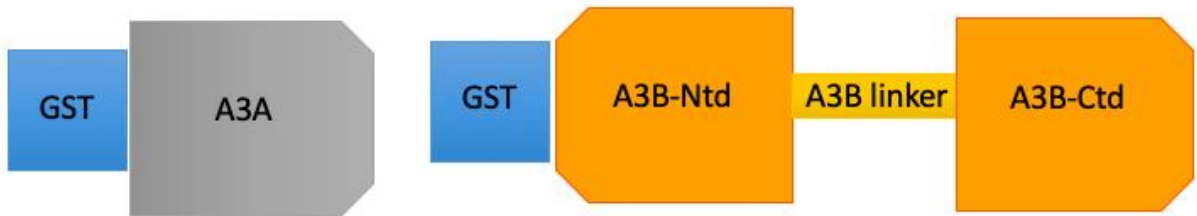
Chapter 2 - Materials and Methods

Enzyme Expression and Purification

Glutathione S-transferase (GST)-tagged AID was incorporated into pGEX5.3 (GE Healthcare, USA). This vector was transformed into DE3 *E. coli*, which were then induced to express the desired protein via the use of 1mM Isopropyl β -D-thiogalactopyranoside (IPTG) at 16°C for 16 hours if colony growth containing the correct sequence took place. In the successful cases, these cells were then lysed in ice-cold 1x phosphate-buffered saline (PBS) using a room temperature French pressure cell press at 1000 psi (Thermospectronic) as per the manufacturer's recommendations (¹¹⁴). The lysate was centrifuged and the supernatant was purified using Glutathione Sepharose high-performance beads (Amersham) as per the manufacturer's instructions. The GST-AID was then stored at -80°C in a solution of 100 mM NaCl, 20 mM Tris-HCl pH 7.5, and 1 mM dithiothreitol (DTT).

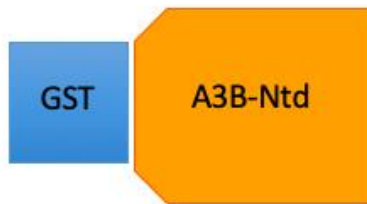
A3A (HMGO Gene Nomenclature Committee [HGNC] name APOBEC3A-003, Vega transcript OTTHUMT00000321238, ¹¹⁵), A3B (GenBank accession no. NM004900, HGNC APOBEC3B-003, Vega transcript OTTHUMP00000199090, ¹¹⁵), Ntd-A3B, Ctd-A3B, a A3B-Ntd A3A-Ctd chimera (BNAC), A3F (HGNC APOBEC3F-001, Vega transcript OTTHUMT00000321216, ^{114,115}), and A3G (HGNC APOBEC3G-001, Vega transcript OTTHUMT00000321219.1, ¹¹⁵) were inserted into GST-containing pcDNA 3.1 (+) vectors (Figures 8 and 9).

Controls:



Experimental Inserts:

Ntd-A3B



A3B-Ctd



BNAC (Ntd-A3B A3A-Ctd Chimera)



Figure 8: Design of Ntd-A3B, Ctd-A3B, and BNAC

As I hypothesized the structural differences of A3A (single domain) and A3B (double domain) were due to domain structure, I designed truncated and chimeric proteins to determine if domain structure was the key to A3B's increased activity on distally damaged substrates. Single domain A3A and double-domain A3B would act as controls. Meanwhile, A3B would be synthesized as its separate domains, the regulatory Ntd-A3B and the catalytic Ctd-A3B. I predict that the presence of a regulatory domain was the reason for A3B's increased activity on distally damaged substrates. Therefore, both Ntd-A3B and Ctd-A3B should display decreased activity relative to normal double domain A3B. The limited literature on the topic suggests that entire domain chimeras are possible, and whichever enzymes composed the Ctd of the chimera dictates its activity. BNAC (A3B Ntd, A3A Ctd) is a chimera designed to provide increased evidence for this theory. As the Ctd is comprised of A3A, yet the regulatory domain of A3B is hypothesized to increase activity on distally damaged substrates, the chimera should have increased activity to relative to normal single domain A3A.

Key:

Random nucleotides for restriction enzyme attachment

~2% of GST that was getting cut off by restriction enzymes while prepping the segment for vector ligation

Linker

Leu, the final amino acid of the Ntd

Stop codons

BamHI restriction site

EcoRV restriction site

Met, the first amino acid of the Ctd

Intron

Silent aspartate mutation to prevent the creation of an additional EcoRV restriction site

Naturally occurring linker region between A3B domains

A3B-Ntd

AGC TAC CGG ATC CCC AGG AAC TCG GAC ACT CTG GAC ACC ACT ATG
AAT CCA CAG ATC AGA AAT CCG ATG GAG CGG ATG TAT CGA GAC
ACA TTC TAC GAC AAC TTT GAA AAC GAA CCC ATC CTC TAT GGT CGG
AGC TAC ACT TGG CTG TGC TAT GAA GTG AAA ATA AAG AGG GGC
CGC TCA AAT CTC CTT TGG GAC ACA GGG GTC TTT CGA GGC CAG
GTG TAT TTC AAG CCT CAG TAC CAC GCA GAA ATG TGC TTC CTC TCT
TGG TTC TGT GGC AAC CAG CTG CCT GCT TAC AAG TGT TTC CAG ATC
ACC TGG TTT GTA TCC TGG ACC CCC TGC CCG GAC TGT GTG GCG
AAG CTG GCC GAA TTC CTG TCT GAG CAC CCC AAT GTC ACC CTG ACC
ATC TCT GCC GCC CGC CTC TAC TAC TAC TGG GAA AGA GAT TAC CGA
AGG GCG CTC TGC AGG CTG AGT CAG GCA GGA GCC CGC GTG ACG
ATC ATG GAC TAT GAA GAA TTT GCA TAC TGC TGG GAA AAC TTT GTG
TAC AAT GAA GGT CAG CAA TTC ATG CCT TGG TAC AAA TTC GAT GAA
AAT TAT GCA TTC CTG CAC CGC ACG CTA AAG GAG ATT CTC AGA TAC
CTG TAA TGA TAA GAT ATC GCA GTG

A3B-Ctd

AGC TAC CGG ATC CCC AGG AAC TCG GAC ACT TCG GAC ACC ACT ATG
GAC CCA GAC ACA TTC ACT TTC AAC TTT AAT AAT GAC CCT TTG GTC
CTT CGA CGG CGC CAG ACC TAC TTG TGC TAT GAG GTG GAG CGC
CTG GAC AAT GGC ACC TGG GTC CTG ATG GAC CAG CAC ATG GGC
TTT CTA TGC AAC GAG GCT AAG AAT CTT CTC TGT GGC TTT TAC GGC
CGC CAT GCG GAG CTG CGC TTC TTG GAC CTG GTT CCT TCT TTG CAG

TTG GAC CCG GCC CAG ATC TAC AGG GTC ACT TGG TTC ATC TCC TGG
 AGC CCC TGC TTC TCC TGG GGC TGT GCC GGG GAA GTG CGT GCG
 TTC CTT CAG GAG AAC ACA CAC GTG AGA CTG CGC ATC TTC GCT GCC
 CGC ATC TAT GAT TAC GAC CCC CTA TAT AAG GAG GCG CTG CAA ATG
 CTG CGG GAT GCT GGG GCC CAA GTC TCC ATC ATG ACC TAC GAT
 GAG TTT GAG TAC TGC TGG GAC ACC TTT GTG TAC CGC CAG GGA
 TGT CCC TTC CAG CCC TGG GAT GGA CTA GAG GAG CAC AGC CAA
 GCC CTG AGT GGG AGG CTG CGG GCC ATT CTC CAG AAT CAG GGA
 AAC TGA TAA TAA GAT ATC GCA GTG

A3B Ntd-A3A Ctd (BNAC)

AGC TAC CCG ATC CCG AGG AAC TCG GAC ACT CTG GAC ACC ACT ATG
 AAT CCA CAG ATC AGA AAT CCG ATG GAG CGG ATG TAT CGA GAC
 ACA TTC TAC GAC AAC TTT GAA AAC GAA CCC ATC CTC TAT GGT CGG
 AGC TAC ACT TGG CTG TGC TAT GAA GTG AAA ATA AAG AGG GGC
 CGC TCA AAT CTC CTT TGG GAC ACA GGG GTC TTT CGA GGC CAG
 GTG TAT TTC AAG CCT CAG TAC CAC GCA GAA ATG TGC TTC CTC TCT
 TGG TTC TGT GGC AAC CAG CTG CCT GCT TAC AAG TGT TTC CAG ATC
 ACC TGG TTT GTA TCC TGG ACC CCC TGC CCG GAC TGT GTG GCG
 AAG CTG GCC GAA TTC CTG TCT GAG CAC CCC AAT GTC ACC CTG ACC
 ATC TCT GCC GCC CGC CTC TAC TAC TAC TGG GAA AGA GAT TAC CGA
 AGG GCG CTC TGC AGG CTG AGT CAG GCA GGA GCC CGC GTG ACG
 ATC ATG GAC TAT GAA GAA TTT GCA TAC TGC TGG GAA AAC TTT GTG
 TAC AAT GAA GGT CAG CAA TTC ATG CCT TGG TAC AAA TTC GAT GAA
 AAT TAT GCA TTC CTG CAC CGC ACG CTA AAG GAG ATT CTC AGA TAC
 CTG ATG GAA GCC AGC CCA GCA TCC GGG CCC AGA CAC TTG ATG
 GAT CCA CAC ATA TTC ACT TCC AAC TTT AAC AAT GGC ATT GGA AGG
 CAT AAG ACC TAC CTG TGC TAC GAA GTG GAG CGC CTG GAC AAT
 GGC ACC TCG GTC AAG ATG GAC CAG CAC AGG GGC TTT CTA CAC
 AAC CAG GCT AAG AA TCT TCT CTG TGG CTT TTA CGG CCG CCA TGC
 GGA GCT GCG CTT CTT GGA CCT GGT TCC TTC TTT GCA GTT GGA CCC
 GGC CCA GAT CTA CAG GGT CAC TTG GTT CAT CTC CTG GAG CCC CTG
 CTT CTC CTG GGG CTG TGC CGG GGA AGT GCG TGC GTT CCT TCA
 GGA GAA CAC ACA CGT GAG ACT GCG TAT CTT CGC TGC CCG CAT
 CTA TGA TTA CGA CCC CCT ATA TAA GGA GGC ACT GCA AAT GCT GCG
 GGA TGC TGG GGC CCA AGT CTC CAT CAT GAC CTA CGA TG GTA AGA
 ATG GAA GGT TCA GGT GGG GTG GGG TGG GTG GGG GCA GGA GAG GTT
 CCT GGG AAG AAA AGG AGA AAG GCC TTG GTC TGC TGC CTG CAG AAA
 CGA TGG CTG GAC TCT GGG ACC TGA CTT TGG GGT CGA TGG GAA GAG
 AGA GGC CAG GCC AGG AGA TGT GGG CCC AGG GAG GGC AGG GAG AGT
 GGC TGG AAG TGG AAG CAG AAC TTG GGG CTT TCT GAA AGA ATG AGA

ACT GGG CTG GCC CAG ATT CCA ATG GGA AGG AAC TGC CTG ATG AAG
 GAG CTA AGT CCC TAG GGG AGG GAG AGG GAA AGG AGG GAC TGA AAC
 CAG GAT GTG GGA AGT CTG TCC TGA GAG TCA TGG GCC CTA GGT GCC
 ACC CCG ATC CCA CAG CGG GAG CGT GAC TTA TCT CCC CTG TCC CTT TTC
 AGA A TTT AAG CAC TGC TGG GAC ACC TTT GTG GAC CAC CAG GGA
 TGT CCC TTC CAG CCC TGG GAT GGA CTA GAT GAG CAC AGC CAA
 GCC CTG AGT GGG AGG CTG CGG GCC ATT CTC CAG AAT CAG GGA
 AAC TAA TGA TAA GAT ATC GCA GTG

Figure 9: Sequences of Ntd-A3B, Ctd-A3B, and the BNAC chimera designed for insertion into GST-containing pcDNA 3.1/V5-His TOPO vector

The purpose of the truncated and chimeric A3B design was to provide two methods to prove the regulatory Ntd of A3B was responsible for its increased activity on distally damaged ssDNA substrates. The sequence for A3B was separated into its respective domains, the Ntd and the Ctd (each synthesized uniquely). The Ntd of A3B and A3A, the equivalent of an active “Ctd”, were combined to form a chimera (BNAC). If A3B’s Ntd was responsible for the increase in activity on distally damaged substrates, Ctd-A3B should exhibit lower activity than normal-double domain A3B. Alternatively, BNAC should demonstrate increased activity relative to normal A3A, which lacks a regulatory domain. Restriction sites, stop codons, and additional required nucleotide sequences are indicated in the key above. Cloning involved the PCR of these sequences, a BamHI/BamHI restriction digest, an overnight ligation into GST-containing pcDNA 3.1/V5-His TOPO vector, and transformation into XL1-Blue or Top10 cells. Transformed cells were plated onto ampicillin-containing agar and grown overnight. Colonies were picked, and their plasmid DNA isolated to be sent for sequencing at GenScript.

Some GST-A3 open reading frames (ORF) were EcoRV/KpnI-flanked (A3A, A3B, A3F, A3G), and some BamHI/BamHI-flanked (Ntd-A3B, Ctd-A3B, BNAC); A3A and BNAC contained a single native intron from APOBEC3A-003 (Figure 9,¹¹⁴) to prevent cytotoxicity during bacteria-based cloning similar to the methods for GST-AID expression. Without the presence of the intron, the highly active A3A appears to cause a level of DNA damage that the bacteria cannot recover from; the presence of an intron effectively renders A3A inactive in prokaryotes because they lack the splicing mechanisms necessary to activate the protein (¹¹⁶⁻¹²⁵). This intron did not affect the final product because the eukaryotic cell line utilized, HEK-293T cells, were capable of

removing the region (¹¹⁶⁻¹²⁵). Fifty 10 cm plates, each seeded with 5×10^5 HEK-293T cells, were transfected with 5 μ g of expression plasmid per plate using Polyjet transfection reagent (FroggaBio). Cells were incubated for 48 hours, collected, and resuspended in 500 mM NaCl, 50 mM Phosphate Buffer pH 8.2, 0.2 mM PMSF, and 50 μ g/ml RNase A. A3A/B/F/G, Ntd-A3B, Ctd-A3B, and BNAC expression were confirmed by western blotting using an anti-GST primary (Abcam, catalogue number ab9085) at a 1:500 dilution in 1x PBS followed by secondary Goat anti-Rabbit IgG (Santa Cruz Biotechnology, catalogue number sc-2004) diluted to 1:500 in 1x PBS as per the manufacturer's instructions. For the western blot, proteins were transferred to a nitrocellulose membrane and blocked with 5% skim milk at room temperature for one hour, then incubated with the primary antibody while shaking at room temperature overnight. The membrane was then washed three times with TBST at room temperature, with each wash lasting five minutes. The membrane was then incubated with the secondary antibody in 5% skim milk at room temperature for one hour. Protein concentrations were then assessed using standard image scanning methods.

Following the confirmation of protein presence on the western blot, cells were then lysed in ice-cold 1x PBS, again using a French pressure cell press at room temperature and 1000 psi as per the manufacturer's recommendations (Thermospectronic). The lysate was cleared by ultracentrifugation and the supernatant was purified using Glutathione Sepharose high-performance beads (Amersham) as per the manufacturer's instructions. All GST-tagged enzymes were stored at -80°C in 100 mM NaCl, 20mM Tris-HCl pH 7.5, 1mM DTT, 5% Glycerol, and 50 μ g/ml BSA. Yields of purified enzymes were assessed on Coomassie-stained SDS-PAGE (Figure 10), and protein concentrations were

determined using bovine serum albumin (BSA) standard curves, as previously described^(126,127).

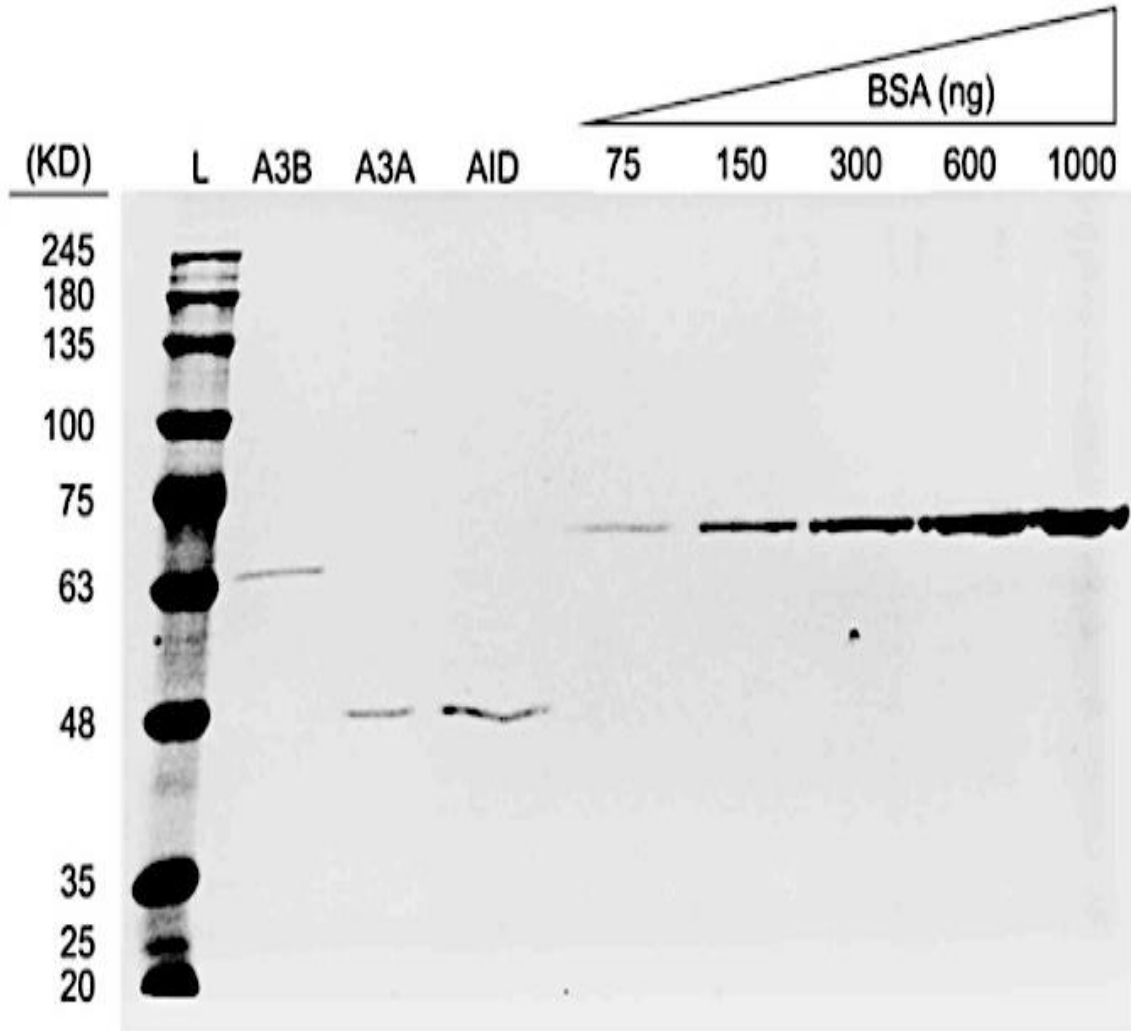


Figure 10: SDS Gel to Verify Purity of GST-A3A and GST-A3B Utilized for Experimentation

GST-A3B, GST-A3A, and human GST-AID preparations (n=2) were quantified using a series of BSA standards. Purity can be determined by little to no additional bands being present in each lane, and size can be determined by referencing a standardized ladder (see lane labelled L).

***Reprinted from AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. 1863/11, Diamond CP¹, Im J¹, Button EA¹, Huebert DNG¹, King JJ¹, Borzooee F¹, Abdouni HS¹, Bacque L¹, McCarthy E¹, Fifield H¹, Berghuis LM¹, Larijani M². AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation. Copyright 2019, with permission from Elsevier.*

GST-A3A and GST-A3B, in addition to all other enzymes synthesized, were stored at -80°C in 100 mM NaCl, 100mM NaCl, 20mM Tris-HCl pH 7.5, and 1 mM DTT.

To date, Ctd-A3B and BNAC have not produced colonies containing the correct sequence, and have therefore not been expressed in HEK-293T cells.

Preparation of Substrates and the Alkaline Cleavage Enzyme Assay

To complete the activity trials of A3A and A3B on undamaged ssDNA substrates and substrates containing an 8oxoG in the -1, -7, +7, and both -7 and +7 positions relative to the dC, alkaline cleavage using 5'-[γ -P32] labelled substrates was employed. Substrates were synthesized and fast protein liquid chromatography (FPLC)-purified by Integrated DNA Technologies, Midland, or Trilink. Oligonucleotide of a 2.5 pmol concentration was 5'-labelled using [γ -P32] dATP with the aid of T4 polynucleotide kinase (New England Biolabs). Once labelled, the oligonucleotides were purified using mini Quick Spin Oligo Columns (Roche), diluted with 1x Tris-Ethylenediaminetetraacetic acid (TE) buffer pH 8 to 50 fmol, and stored at -20°C. The alkaline cleavage assay, previously described in (¹²⁸⁻¹³²), was utilized to verify substrate size with an enzyme preparation known to be active. The alkaline cleavage assay was then utilized to compare the activity of A3A and A3B on undamaged ssDNA substrates and substrates containing an 8oxoG in the -1, -7, +7, and both -7 and +7 positions relative to the dC. This assay used 1 μ l substrate at a substrate concentration ranging from 0.1-10 nM, 6 μ l of reaction buffer (25 mM HEPES buffer pH

5.5), and 3 μ l of enzyme (2-10 ng). The total reaction volumes were 10 μ l. The concentrations of Ntd-A3B, Ctd-A3B, BNAC, A3F, and A3G are unknown at this time because the proteins have yet to be synthesized.

For initial determination of enzyme activity on normal versus damaged substrates, substrate concentrations of 7.5, 5, 4, 2.5, 1.25, 0.625, 0.315, and 0.15 nM were used. For measurement of initial deamination velocity kinetics, a range of substrate concentrations including 7.5, 5, 4, 2.5, 1.25, 0.625, 0.315, and 0.15 nM were assessed. Reactions were incubated at 37°C for 1 hour for A3A and overnight (8 hours) for A3B, followed by heat inactivation of the APOBEC3 enzyme via incubation at 80°C for 20 minutes. The reaction volume was then doubled to 20 μ l by the addition of Uracil-DNA Glycosylase (UDG) (1 μ l containing 1 unit), 2 μ l 10x UDG buffer (New England Biolabs), and 7 μ l ddH₂O and incubated at 37°C for 30 minutes to excise deaminated dCs (now deoxyuridines, dUs). Finally, 100 mM NaOH was added to the reaction mixture which was then incubated at 96°C for 5 minutes to cleave the abasic site left as a result of dU excision. Each sample was loaded on 14-16% denaturing polyacrylamide gels for electrophoresis. Gels were exposed to a Kodak Storage Phosphor Screen and visualized using a phosphorimager (Bio-Rad) (Figures 11 and 12).

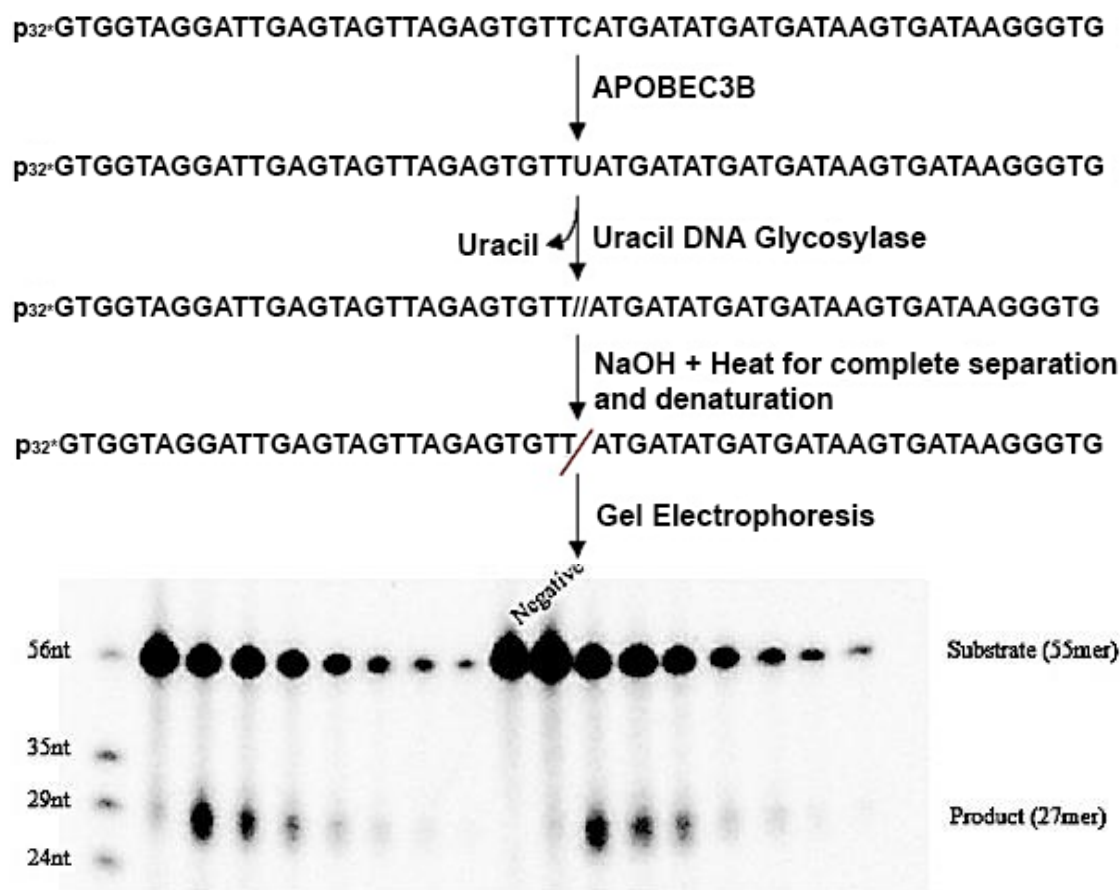


Figure 12: Model Explaining the Procedure of Alkaline Cleavage

Alkaline cleavage featured heavily in this work. The procedure begins with radioactively labelling ssDNA substrates using 5'-[γ -P32]. These substrates are then exposed to the desired APOBEC3 enzyme, represented in this diagram as APOBEC3B. dCs will be converted to dU during this reaction. After an appropriate incubation time at 37 °C, Uracil DNA Glycosylase is added to excise the dU now present in the DNA. This leaves an abasic site, which results in cleavage of the substrate into 2 parts after exposure to NaOH and heat. The reaction mixture, which is run on a 14-16% acrylamide gel in the green gel tank pictured above, will separate into distinct substrate and product bands. The substrate band represents uncleaved substrate, and the product band represents the substrate size after cleavage of the known dC site. The other portion of the substrate after the dC is not visible on the gel because the radioactive tag is only present on the 5' end.

****Figure used with the permission of Junbum Im and Cody Diamond. © 2018 Junbum Im and Cody Diamond.**

The comparison of A3A, A3B, Ntd-A3B, Ctd-A3B, and a chimera of Ntd-A3B and A3A (BNAC) activity on undamaged ssDNA substrates and substrates containing an 8oxoG in the -1, -7, +7, and both -7 and +7 positions relative to the dC will be carried out in an identical matter (Figures 11 and 12). Longer incubation times may be required for Ntd-A3B, Ctd-A3B, and BNAC, but these times are unable to be determined until these enzymes have been synthesized.

Trials assessing A3A, A3B, Ntd-A3B, Ctd-A3B, BNAC, A3F, Ntd-A3F, Ctd-A3F, A3G, Ntd-A3G, and Ctd-A3G activity upon exposure to 8oxoG (-1, -7, +7, -7 and +7), 8oxoA, O6MeG, O4MeT, and 1MeA-containing ssDNA substrates relative to a control undamaged substrate will also utilize the same approach (Figures 11 and 12).

Design of Ntd-A3B, Ctd-A3B, and BNAC (Ntd-A3B, A3A as the “Ctd”) Chimera Sequences

In order to determine if the presence of A3B's regulatory (N-terminal) domain was the reason for its increased activity on distally damaged substrates, two truncated and one chimeric versions of A3B were designed. When designing the hybrid enzymes for cloning, several aspects had to be considered for the DNA sequences. Firstly, as the vector I intended to use for cloning, pcDNA 3.1 (+) Mammalian Expression Vector (ThermoFisher Scientific, catalogue number V79020), contained a GST tag-encoding sequence that would be partially cut by the BamHI insert site, each hybrid enzyme I designed had to contain the last 2% of GST's sequence plus a linker sequence specific to the vector sequence. Secondly, the restriction sites I intended to use were 5' BamHI/ 3' EcoRV, so the appropriate cut sequences with random nucleotide overhang had to be

included to form sticky ends that would ligate into the pcDNA 3.1 (+) vector. Thirdly, several stop codons were required at the end of the sequence to minimize any possibility of read-through. As for the control wild-type enzymes, the sequences of A3A and A3B were already present in our lab's stock of GST-containing pcDNA 3.1 (+). A literature review of A3 chimeras and A3B structure revealed that the Ntd and Ctd of A3B could be easily demarcated from one another by the border of methionine-194 (Met194). Therefore, Ntd-A3B contained residues 1 through 193 (which includes A3B's natural domain linker sequence), and Ctd-A3B contained residues 194 through 382 (¹⁰¹). The Ntd-A3B A3A-Ctd chimera I designed (BNAC) contained residues 1 through 193 of A3B and the sequence for A3A (A3A has only a single domain, and is therefore equivalent to a "CD2" active domain). One crucial change was made to the nucleotide base sequence of A3B to eliminate its naturally occurring internal BamHI cut site, but the amino acid sequence of A3B remained the same. IDTs GeneBlocks ordering software also had restrictions on the number of repetitions a sequence could have, necessitating several silent nucleotide base mutations to all the A3B sequences to deal with the logistics of this; the amino acid sequence would of course remain identical to the original (Figure 9).

Originally, the Ntd and Ctd of A3B and the BNAC chimera were to be synthesized separately by ordering these sequences from GeneBlocks (IDT) (Figures 8 and 9). After insertion into the vector, the products were transformed into Top10 and XL1-Blue cells and grown on LB+amp plates overnight at 37 °C. Colonies were then sent for sequencing by GenScript.

Unfortunately, only Ntd-A3B was successfully ligated into GST-containing pcDNA 3.1/V5-His TOPO and resulted in colonies with the correct orientation upon sequencing.

As a result, Ctd-A3B and BNAC were re-ordered, this time in a vector rather than as a DNA segment. Polymerase chain reaction (PCR) of BNAC became possible for the first time, and there was some evidence on agarose gels that cloning may have been successful. However, GenScript was unable to produce a complete read of the sequence so correct insertion and orientation were impossible to determine. Ctd-A3B did not produce colonies, and will require further troubleshooting.

PCR and Cloning of Hybrid Chimeric A3A/B and Truncated A3B Enzymes

The DNA fragments ordered from GeneBlocks (IDT) were of the following sizes: Ntd-A3B was 639 bp, Ctd-A3B was 633 bp, and BNAC was 1652 bp. As the stock shipped in small amounts, PCR amplification of the fragments was a necessary step. 50 µl reactions were utilized, with the reaction mixture containing 1 µl of 1 ng/µl of DNA, 5 µl of 10x PCR buffer, 10mM of deoxynucleoside triphosphate (dNTP), 10 µM of forward and reverse primers, Taq DNA polymerase (New England BioLabs), and dH₂O as per the manufacturer's instructions. The primers utilized are listed below; they initially contained the complementary sequence for a 3' EcoRV cut site, but new primers (specifically those listed below) containing a 3' BamHI site were ordered soon afterwards and utilized throughout the majority of the PCR trials.

Forward primer for Ntd-A3B, Ctd-A3B, and BNAC:

5' – GGATCCCCAGGAACTCGGACACTCTGGACACCACTATG – 3'

Reverse primer for Ntd-A3B:

5' - GCGATGGGATCCTTATCATTACAGGTATCTGAGAATCTCCTTTAGCG - 3'

Reverse primer for Ctd-A3B:

5' - GCGATGGGATCCTTATTATCAGTTTCCCTGATTCTGGAGAATGGCCC - 3'

Reverse primer for BNAC:

5' - CACTGCGGGATCCTTATCATTATCAATTTCTTGATTTTGGAGGATGG
CTCG - 3'

Samples were incubated in a PCR machine at 94°C for 2 minutes, then underwent 30 cycles of 94°C for 30 seconds + 56°C for 30 seconds + 72°C for 1 minute, then finally a 72°C periods for 10 minutes followed by a hold at 4°C. The sizes of the amplified fragments were then verified on a 1x Tris-borate-Ethylenediaminetetraacetic acid (TBE) 0.9% agarose gel (if not performing a gel extraction) or a 0.5x Tris-acetate-Ethylenediaminetetraacetic acid (TAE) 0.9% agarose gel (if performing a gel extraction) run at an average of 120 V for a minimum of 1 hour. A PCR clean-up was then performed using a QIAquick PCR Purification Kit (Qiagen) as per the manufacturer's instructions with the exception of using two runs through the elution column rather than one and pre-warming the elution buffer in a 37°C water bath. The product concentration was then measured using a NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific) as per the manufacturer's instructions.

Next, ligation of the gene fragments into pcDNA 3.1 (+) was attempted. The ligation reaction consisted of 10 µl total volume reactions, utilizing DNA fragments diluted to 30 ng/µl. The reactions contained 5ng of DNA fragment (Ntd-A3B, Ctd-A3B, and BNAC, respectively), 1 µl of pcDNA 3.1 (+), T4 DNA ligase (New England BioLabs, catalogue

number M0202S), and dH₂O as per the manufacturer's instructions. The ligation reactions were then incubated at 16°C for 16 hours followed by 1 hour at 25°C. The newly ligated vectors were then transformed into competent cells; XL1-Blues and One Shot™ Top10 Chemically Competent *E. coli* (ThermoFisher Scientific catalogue number C404003) were both utilized during the troubleshooting period. The entire 10 µl ligation reaction was added to 40 µl of competent cells, and then stored on ice for 30 minutes. The competent cells were then heat-shocked in a 42°C water bath for 60 seconds before being returned to ice for a further 60 seconds. 150 µl of the transformation reaction was then plated onto lysogeny broth (LB)/ 50 µl/ml ampicillin plates pre-warmed at 37°C. The plates were then incubated at 37°C overnight (approximately 16 hours).

The following day colonies were picked using a disposable pipette tip, then grown overnight (approximately 16 hours) in 5 mL LB broth+10 µl of 50 µl/ml ampicillin at 37°C with shaking incubator at 225 cycles/minute. The reaction tubes were then centrifuged at 3000rpm for 10 minutes and their LB/ampicillin broth decanted. The DNA contained within the pellet was then extracted using a QIAprep Spin Miniprep Kit (Qiagen) as per the manufacturer's instructions. The product concentration was measured using a NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific) as per the manufacturer's instructions and the product's size verified using a 1x TBE 0.9% agarose gel before being sent for sequencing (GenScript). If the product appeared to contain an elevated salt or contaminant content, gel extraction using a QIAquick Gel Extraction Kit (Qiagen) as per the manufacturer's instructions was performed prior to sending the product for sequencing.

Methods for Data Collection and Analysis

As previously described in other papers utilizing the standard alkaline cleavage deamination assay (^{105,126,133,134}), Quantity One 1-D Analysis Software (Bio-Rad) was used to conduct band densitometry of the substrate and product bands. Each experiment was repeated in triplicate with multiple independently purified preparations of A3s where possible. Thus, each data point on enzyme activity graphs represents an average of 6-9 independent values. Michaelis-Menten enzyme activity graphs were generated using GraphPad Prism software, where error bars were used to represent Standard Deviation. K_m and V_{max} values were obtained by the non-linear regression Michaelis-Menten analysis in GraphPad, and used to derive K_{cat} based on the concentration of enzymes used. Agarose gels were photographed under UV light and analyzed using Quantity One.

Chapter 3 - Results

Substrates Containing Distal 8oxoGs Significantly Affect A3B but Not A3A Activity, and Minimally Affect AID Activity

All experiments published in our 2019 manuscript Diamond et al. **AID, APOBEC3A and APOBEC3B efficiently deaminate deoxycytidines neighboring DNA damage induced by oxidation or alkylation** which formed the basis for my M.Sc. research, involved the damaged nucleotide being present in the -1 or -2 position relative to the dC. I was interested in examining the effect of damaged nucleotides if they were present more distally than -1 or -2, if they were present upstream of the dC, or if there was more than one lesion in the substrate. This avenue was explored by designing substrates with an 8oxoG in the -7, +7, and both the -7 and +7 positions (Figure 11). This permitted the investigation of the effect of not just distal damage relative to the dC, but also damage upstream and the concept of mutational load (more than one damaged base in the substrate). -7 and +7 were chosen simply because both the undamaged control substrate and the proximally damaged control substrate (-1 8oxoG) coincidentally had a G in both the -7 and +7 positions, and G was the ideal negative for 8oxoG in these damaged base experiments. Furthermore, we reasoned that, as little to no work had been completed in this area, any substrate with damaged located greater than 2 nucleotides away from the target dC would suffice to gain preliminary data on the topic while remaining economical. Trials with substrates containing 8oxoG in a variety of locations relative to the target dC should certainly be carried out in the future (see Chapter 5 – Proposal for Future Work), but would require that new undamaged control sequences be designed and ordered. AID,

A3A, and A3B all exhibited activity on these -7, +7, and -7 & +7 substrates (Figures 13, 14, and 15, Table 2), but differences relative to the undamaged control substrate varied between enzymes.

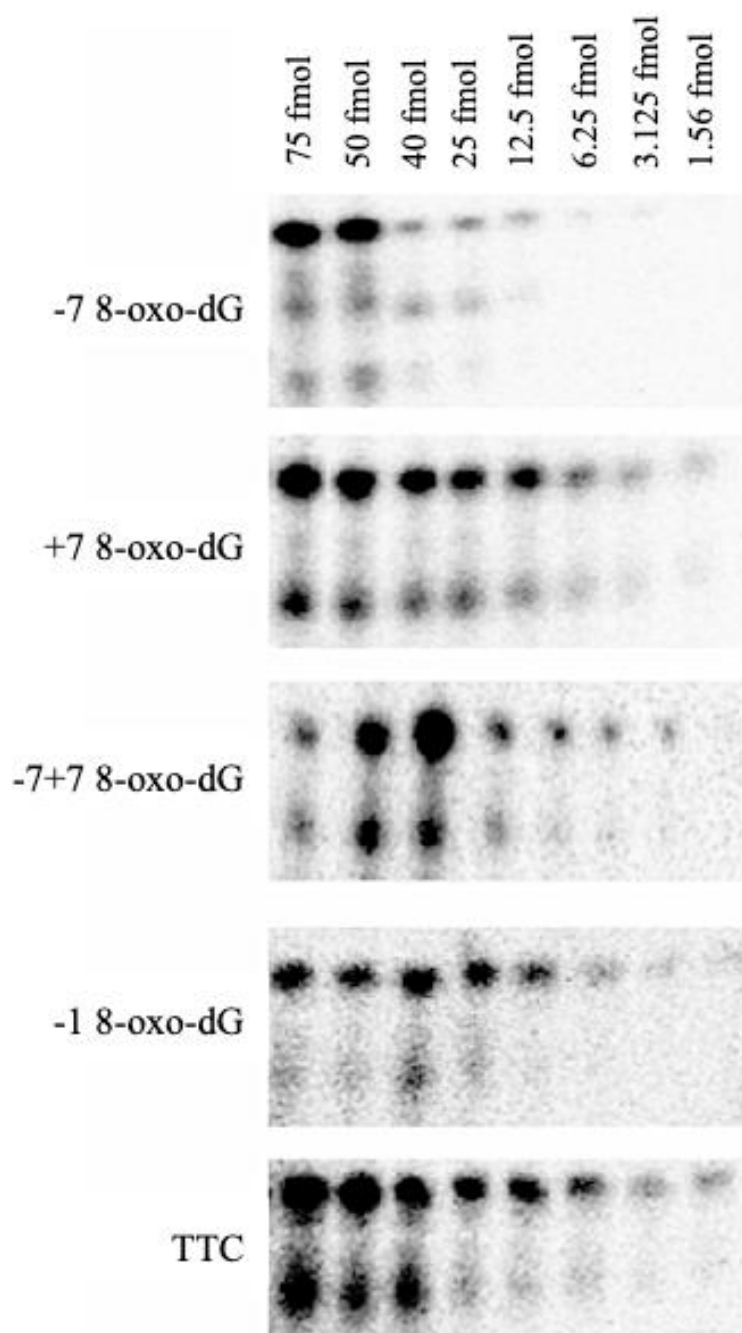


Figure 13: Alkaline Cleavage Gels of AID Acting on ssDNA Substrates Containing 8oxoG in Various Locations and Quantities

As 8oxoG (8-oxo-dG) was found to increase AID activity relative to other damaged substrates, trials to explore the effect of 8oxoG positioning relative to the dC were designed ($n=9$). All substrates were ssDNA, with undamaged TTC and 8-oxo-dG positioned 1 base upstream of the dC acting as controls. Other substrates featured 8-oxo-dG 7 bases upstream, 7 bases downstream, or both.

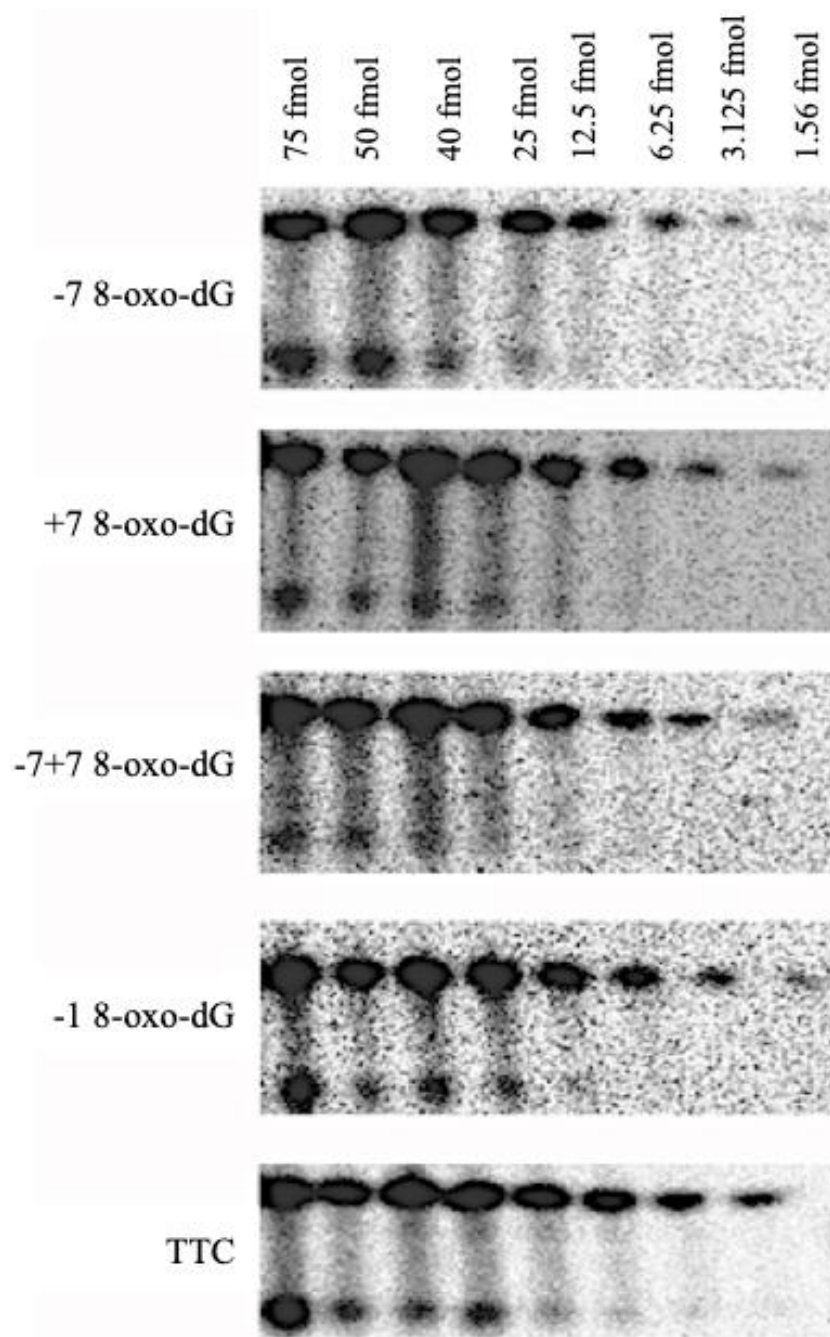


Figure 14: Alkaline Cleavage Gels of A3A Acting on ssDNA Substrates Containing 8oxoG in Various Locations and Quantities

As 8oxoG (8-oxo-dG) was found to increase A3A activity most significantly relative to other damaged substrates, trials to explore the effect of 8oxoG positioning relative to the dC were designed (n=9). All substrates were ssDNA, with undamaged TTC and 8-oxo-dG positioned 1 base upstream of the dC acting as controls. Other substrates featured 8-oxo-dG 7 bases upstream, 7 bases downstream, or both.

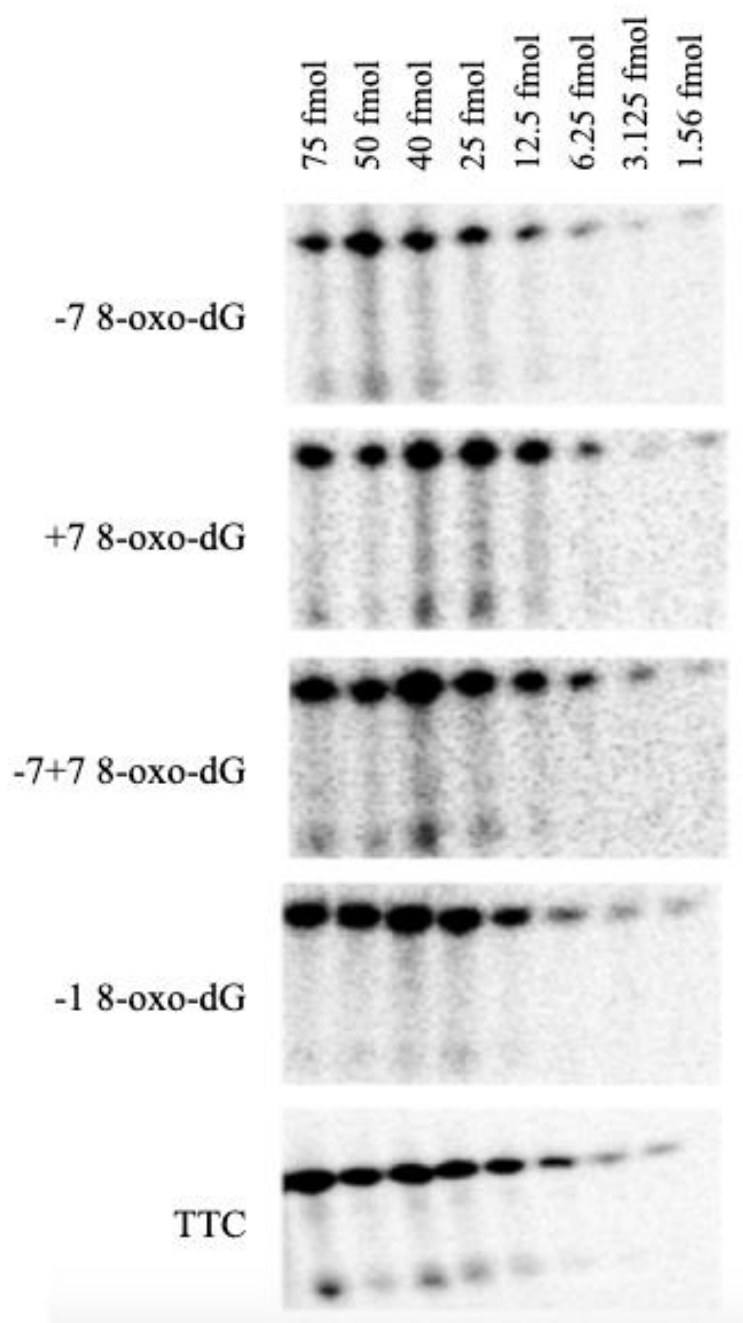


Figure 15: Alkaline Cleavage Gels of A3B Acting on ssDNA Substrates Containing 8oxoG in Various Locations and Quantities

As 8oxoG (8-oxo-dG) was found to increase A3B activity relative to other damaged substrates, trials to explore the effect of 8oxoG positioning relative to the dC were designed ($n=9$). All substrates were ssDNA, with undamaged TTC and 8-oxo-dG positioned 1 base upstream of the dC acting as controls. Other substrates featured 8-oxo-dG 7 bases upstream, 7 bases downstream, or both.

Enzyme	Substrate	K _m (μM)	K _{cat} (min ⁻¹)	K _{cat} /K _m (min ⁻¹ μM ⁻¹)	K _{cat} /K _m fold difference	P value	Significance in a paired t-test of undamaged: damaged substrate
AID	TTC (undamaged)	4.395E2	4.019E-3	9.144E-6			
	T(8oxoG)C (-1 8oxoG)	2.556E15	2.198E10	8.599E-6	9.404E-1	0.2559	ns
	-7 8oxoG	8.387E15	1.126E11	1.342E-5	1.468	0.0280	*
	+7 8oxoG	8.767E15	1.139E11	1.299E-5	1.420	0.0345	*
	-7+7 8oxoG	4.192E2	3.883E-3	9.263E-6	1.013	0.7155	ns
A3A	TTC (undamaged)	4.010E16	6.219E12	1.551E-4			
	T(8oxoG)C (-1 8oxoG)	3.194E16	4.792E12	1.5E-4	9.673E-1	0.5708	ns
	-7 8oxoG	1.271E16	1.577E12	1.241E-4	7.998E-1	0.6683	ns
	+7 8oxoG	3.490E16	5.563E12	1.594E-4	1.027	0.4565	ns
	-7+7 8oxoG	4.596E16	4.952E12	1.077E-4	6.946E-1	0.2992	ns
A3B	TTC (undamaged)	55.93	0.08029	1.4355E-3			
	T(8oxoG)C (-1 8oxoG)	4104	3.376	8.226E-4	5.730E-1	0.8475	ns
	-7 8oxoG	64.62	0.3532	5.466E-3	3.808	0.0122	*
	+7 8oxoG	74.50	0.2828	3.796E-3	2.644	0.0084	**
	-7+7 8oxoG	74.73	0.3419	4.575E-3	3.187	0.0138	*

Table 2: Catalytic parameters of AID, APOBEC3A, and APOBEC3B on distally damaged 8oxoG-containing motifs compared to normal favoured sequence motifs and proximally damaged 8oxoG sequence motifs

AID demonstrated a slight increase in activity on the -7 and +7, substrates (~1.5, ~1.4-fold increase in activity, respectively) (Figures 13 and 16, Table 2).

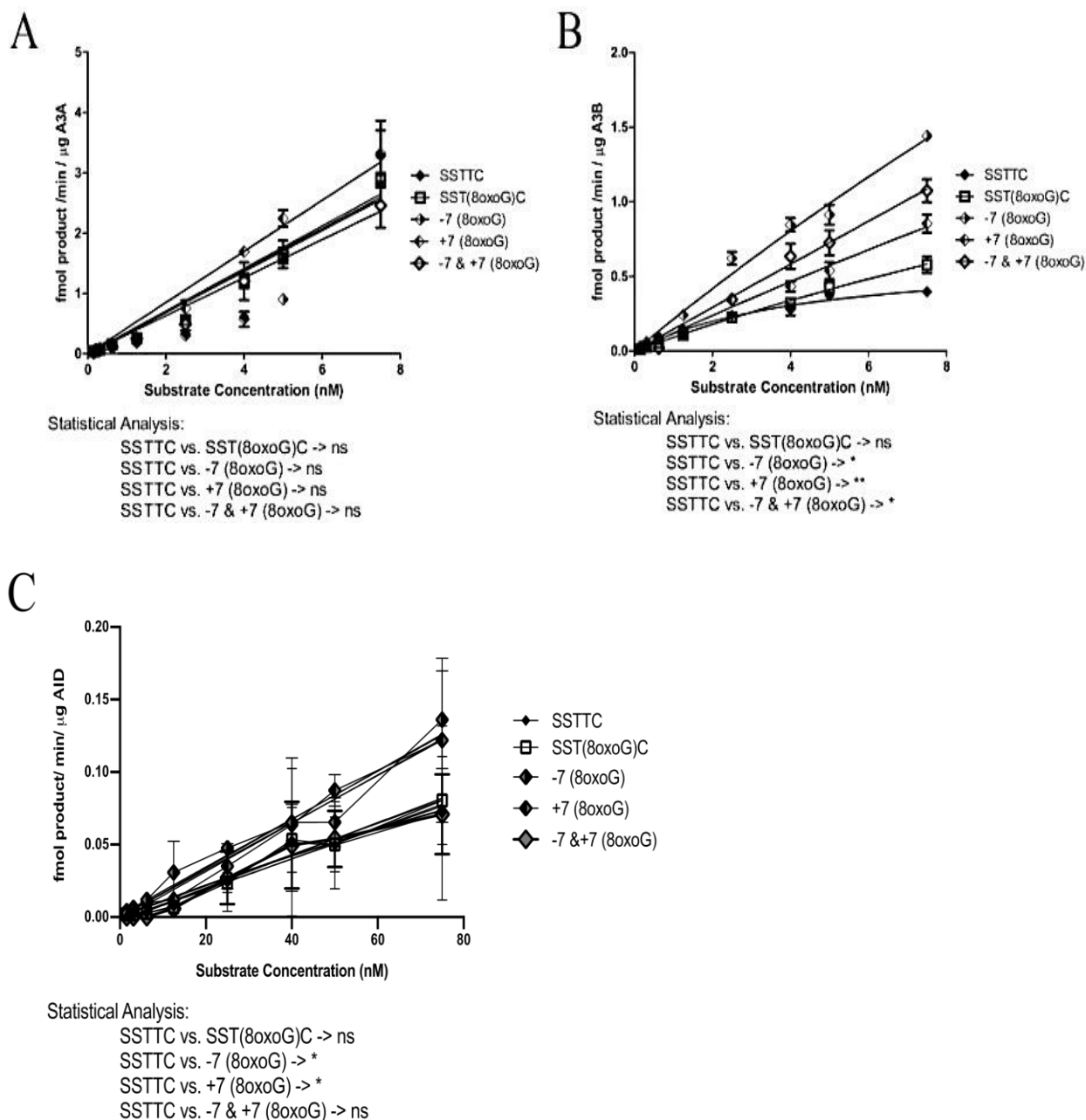


Figure 16: The Effect of 8oxoG Position and Quantity on AID, A3A, and A3B Activity

Michaelis-Menten kinetics graphs of GST-tagged AID, A3A, and A3B activity on ssDNA substrates containing 8oxoG's both distal and proximal to the target dC. These reactions (n=9) contained 1 μ l of substrate at a substrate concentration ranging from 0.1-10 nM, 6 μ l reaction buffer (25 mM HEPES buffer pH 5.5), and 3 μ l enzyme (2-10 ng). The total reaction volumes were 10 μ l. Incubation times varied for each enzyme ranging from 30-240 minutes. The following kinetics results were noted:

- a) Varying the position of 8oxoGs to -7 and/or +7 did not produce significant activity changes for A3A when compared to the control undamaged TTC and -1 (8oxoG)-containing substrates.*
- b) -7, +7, and -7/+7 were found to significantly increase A3B activity relative to -1 8oxoG and the undamaged TTC substrate.*
- c) When AID was exposed to either -7 or +7 (8oxoG) substrates it exhibited higher activity than for native TTC and -1 (8oxoG)-containing substrates.*

A3A was not found to be significantly more active on -7, +7, and -7/+7 substrates than it was on single-stranded TTC (ssTTC) substrates ($p>0.05$) (Figures 14 and 16, Table 2).

However, A3B was found to be ~4 times more active on the -7 8oxoG substrate ($p\leq 0.05$), ~2.6 times more active on +7 substrate ($p\leq 0.01$), and ~3 times more active on the -7/+7 substrate ($p\leq 0.05$) when compared to undamaged substrate values (Figures 15 and 16, Table 2). This result was interesting, as A3A and A3B have approximately 85% sequence homology; their main contrast lies in A3A (and AID) possessing a single domain and A3B having 2 CD domains (¹⁰⁰).

Design of Double Domain Damage Experiments

As A3B (double-domain) (Figures 15 and 16, Table 2) seemed to have an increased activity on distally damaged substrates relative to A3A (single-domain) (Figures 14 and 16, Table 2), I postulated that the regulatory domain (Ntd, or CD1) of double-domain A3s might play a role in the binding of damaged bases.

To test this theory, synthesizing separate domains, the Ntd and the Ctd of A3B, in addition to adding a regulatory domain to A3A (BNAC chimera) seemed appropriate. If my theory that the regulatory Ntd is responsible for increased A3 activity in distally damaged DNA bases is correct, A3B will show greater activity than A3A, Ntd-A3B, and Ctd-A3B on these substrates. Furthermore, adding a regulatory Ntd to A3A should increase its activity beyond that of ordinary single-domain A3A. Ergo, this theory could be proven in two alternate ways, which would strengthen any claim made by the study. The next stage would logically necessitate the study of other double-domain A3s such as A3F and A3G.

In the planned activity trials (Figures 8, 11, 12), A3A, A3B, Ntd-A3B, Ctd-A3B, and an A3B-Ntd A3A-Ctd (BNAC) chimera would be exposed to undamaged TTC ssDNA substrate, TTC substrate containing an 8oxoG in the -1 position, TTC substrate containing an 8oxoG in the -7 position, TTC substrate containing an 8oxoG in the +7 position, and TTC substrate containing 8oxoGs in the -7 and +7 positions. If my hypothesis that the presence of A3B's regulatory (N-terminal) domain was the reason for its increased activity on distally damaged substrates was correct, certain results can be expected (Figure 17).

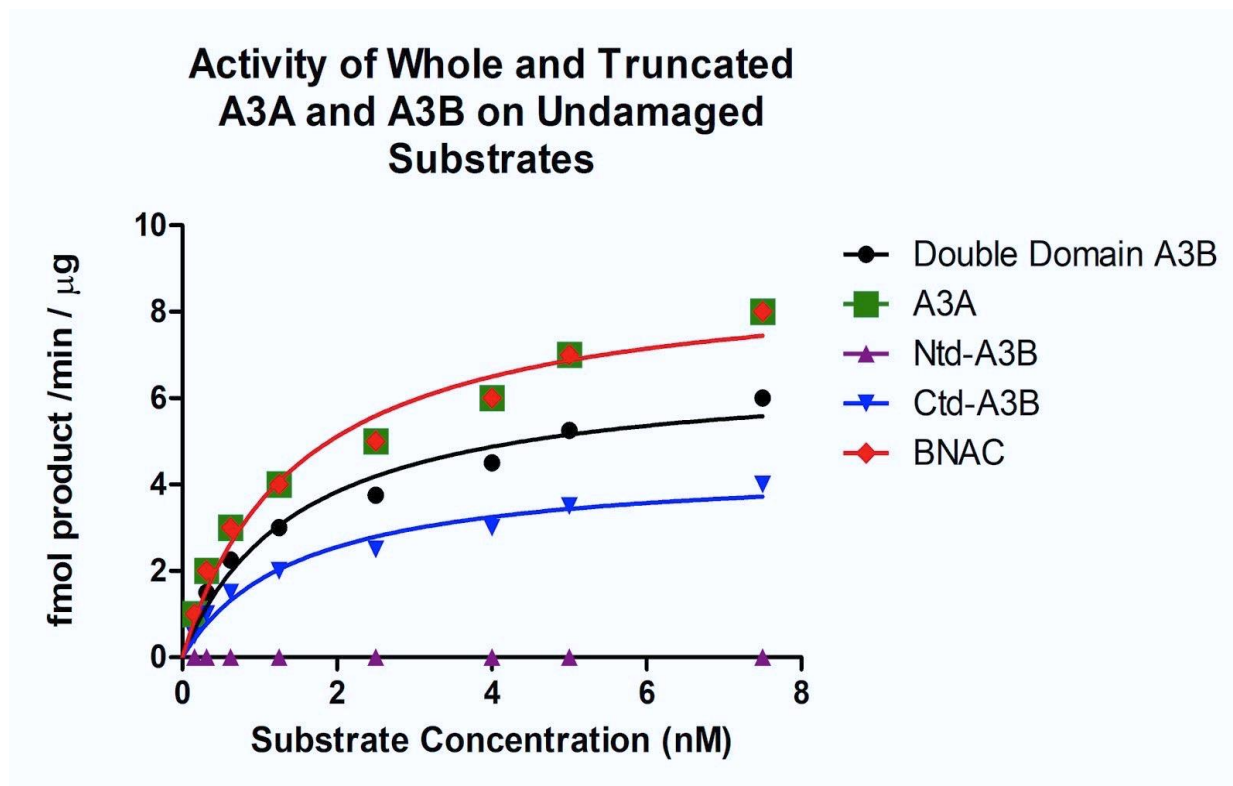


Figure 17: Predicted results of A3A/B double domain activity on undamaged ssDNA substrates based on previous APOBEC3 chimera studies

The numerical values represented here are arbitrary, as difficulties synthesizing Ctd-A3B and BNAC have prevented these alkaline cleavage experiments from being performed at this time. However, this predicted trend is based on the conclusions of other APOBEC3 chimera studies, where the Ctd identity predicts enzymatic activity. Ergo, when A3A is the Ctd of the chimera BNAC, chimeric activity should be similar to that of A3A.

Assuming the substrate was undamaged, wild-type A3B is overall not as active as wild-type A3A, as shown in previous trials (Figures 3, 4, 7, 14, and 15). Ntd-A3B only plays a regulatory role, and would be expected to exhibit limited catalytic activity by itself⁽²⁵⁾. If A3B relies on its Ntd to increase its overall activity, Ctd-A3B alone should be expected to be less active than full (wild-type) double-domain A3B. Based on previous chimera research^(4,5,26,28-32), whichever enzyme is in the Ctd defines overall activity. Therefore, an

A3B chimera containing A3A as its Ctd should behave with similar activity to wild-type A3A, but activity should be equal or increased beyond that of wild-type A3A due to the presence of a regulatory domain that A3A normally lacks.

The above offers two ways in which to prove that the presence of a regulatory domain in double-domain A3s influences their activity on distally damaged bases. If double-domain A3B has increased activity relative to either of its separated domains alone, this could indicate the importance of the regulatory Ntd. Secondly, if the addition of a more active Ctd to A3B's Ntd increases activity beyond that of double-domain A3B (or, alternatively, if the addition of a regulatory Ntd domain to normally single-domain A3A increases the activity of BNAC beyond that of wild-type A3A), this could indicate that the Ntd of A3B plays an important structural role.

Generation of Hybrid Chimeric A3A/B and Truncated A3B Enzymes

Once the ordered sequences encoding for the open reading frames of the hybrid chimeric and truncated enzymes (Figure 9) arrived, the process of cloning these into our expression vectors began. First, the sequences were electrophoresed on agarose gels to verify their size followed by PCR amplification. The initial PCR utilized featured a temperature gradient for optimization. PCR of Ntd-A3B and Ctd-A3B was successful, but PCR of BNAC proved challenging due to its size. Ntd-A3B and Ctd-A3B were shown to be the correct size on an agarose gel (Figure 18).

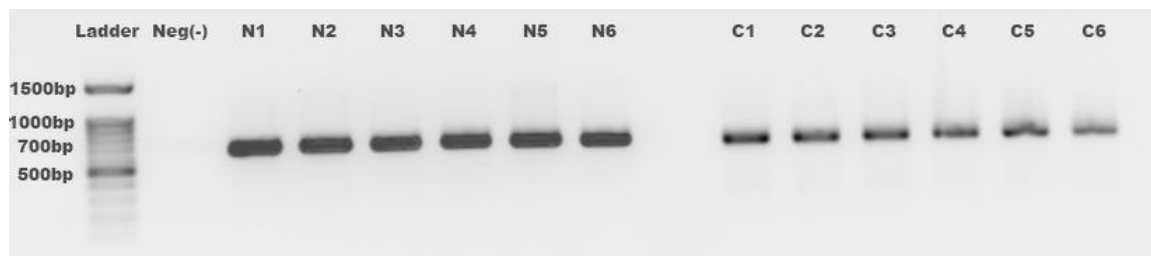


Figure 18: Agarose gel verifying the size of *Ntd-A3B* and *Ctd-A3B* gene fragments post-initial PCR

0.5x TAE 0.9% agarose gel stained with 10 μ l of SYBR Safe and run at 100V for 1 hour and 30 minutes demonstrating the correct sizes of *Ntd-A3B* and *Ctd-A3B*, 639bp and 633bp respectively. This gel establishes that PCR of the *Ntd-A3B* and *Ctd-A3B* gene fragments was possible. BNAC was excluded in this image because several agarose gels run prior to this series of experiments failed to yield any PCR product. N denotes *Ntd-A3B*, with N1 indicating sample 1 of *Ntd-A3B* (n=6 for this PCR trial). C denotes *Ctd-A3B*, with C1 indicating sample 1 of *Ctd-A3B* (n=6 for this PCR trial). Neg (-) indicates a negative that contained no DNA (neither *Ntd-A3B* or *Ctd-A3B*). Ladder denotes a 100bp DNA Ladder (New England Biolabs).

A variety of PCR conditions were planned for further attempts to amplify BNAC but ultimately the BNAC PCR was only successful when I later ordered it in a vector (not as a stand-alone linear sequence as the others described above). When BNAC was ordered in a vector, I also included A3A's native intron in the sequence; this would have made the protein product less toxic to bacteria during cloning competent cells, but the intron would have been easily removed by the processing machinery of 293T cells in the later stages of expression and protein synthesis. This is a common strategy that has been used in the field to circumvent the toxicity of AID/A3 enzymes (¹¹⁶) even as they are expressed in low amounts due to leakiness of mammalian promoters leading to some protein production in bacteria used to propagate the vector during its cloning and to generate sufficient plasmid vector for transfection into mammalian cells (¹¹⁷⁻¹²⁵). As for *Ntd-A3B* and *Ctd-A3B*, the next steps included a BamHI/BamHI restriction digest (the initial

EcoRV site was changed to a second BamHI site via PCR later as part of troubleshooting), an overnight ligation into GST-containing pcDNA 3.1 (+), transformation into competent cells, and overnight growth on ampicillin-containing agar plates. Colonies eventually grew on the Ntd-A3B plates, and these were prepared for sequencing (GenScript) through use of a QIAquick Spin Miniprep Kit (Qiagen). Several colonies contained the correct Ntd-A3B sequence in the correct orientation. Ctd-A3B BamHI/BamHI digest, ligation, transformation, and colony growth were then troubleshooted for several months until we decided to order Ctd-A3B in a vector along with BNAC in a vector. Ctd-A3B in a vector worked for all procedural steps up to but not including transformation/colony growth. Many repetitions and troubleshooting attempts were carried out to try and induce Ctd-A3B colony growth, to no avail (see Discussion).

To summarize these efforts, despite several months of attempts, I was unable to generate expression constructs for the hybrid enzymes necessary to test my second hypothesis. Such expression difficulties are common in the AID/APOBEC field, because the enzymes are quite toxic even at low concentrations (^{124, 125, 135}). Many labs in the field have issues generating expression constructs, often only obtaining catalytically dead versions after sequencing, because any vectors encoding active enzymes prove lethal to bacteria (^{124, 125, 135}). It is worth noting that A3A and A3B are the two most robust and mutagenic (highest catalytic rates) enzymes in the AID/APOBEC family, and this could have amplified the cloning problem described (^{124, 125, 136}).

Efforts are ongoing in the laboratory to solve the cloning issues. Despite the inability to proceed with actual testing of my second hypothesis, below I present hypothetical results

of the outcomes I would expect to obtain from the intended experiment (Figure 17), if my second hypothesis is correct.

Predicted Hypothetical Results of A3A/B Domain Chimeras

In the planned activity trials (Figures 8 and 11), A3A, A3B, Ntd-A3B, Ctd-A3B, and an A3B-Ntd A3A-Ctd (BNAC) chimera would be exposed to undamaged TTC ssDNA substrate, TTC substrate containing an 8oxoG in the -1 position, TTC substrate containing an 8oxoG in the -7 position, TTC substrate containing an 8oxoG in the +7 position, and TTC substrate containing 8oxoGs in the -7 and +7 positions. If my hypothesis that the presence of A3B's regulatory (N-terminal) domain was the reason for its increased activity on distally damaged substrates was correct, certain results can be expected (Figure 17). Assuming the substrate was undamaged, wild-type A3B is overall not as active as wild-type A3A, as shown in previous trials (Figures 3, 4, 7, 14, and 15). Ntd-A3B only plays a regulatory role, and would be expected to exhibit limited catalytic activity by itself ⁽²⁵⁾. If A3B relies on its Ntd to increase its overall activity, Ctd-A3B alone should be expected to be less active than full (wild-type) double-domain A3B. Based on previous chimera research ^(4,5,26,28-32), whichever enzyme is in the Ctd defines overall activity. Therefore, an A3B chimera containing A3A as its Ctd should behave with similar activity to wild-type A3A, but activity should be equal or increased beyond that of wild-type A3A due to the presence of a regulatory domain that A3A normally lacks.

The above offers two ways to examine whether the presence of a regulatory domain in double-domain A3s influences their activity on distally damaged bases. If double-domain A3B has increased activity relative to either of its separated domains alone, this could indicate the importance of the regulatory Ntd. Secondly, if the addition of a more active Ctd to A3B's Ntd increases activity beyond that of double-domain A3B (or, alternatively, if the addition of a regulatory Ntd domain to normally single-domain A3A increases the activity of BNAC beyond that of wild-type A3A), this could indicate that the Ntd of A3B plays an important structural role.

Chapter 4 - Discussion

Discussion

This work offers an expansion of knowledge in the field of AID/A3s by completing the first AID/A3A/A3B activity trials on substrates containing damaged nucleotide bases. Not only did AID, A3A, and A3B prove to be active on DNA substrates containing 8oxoA, 8oxoG, O4MeT, 1MeA, and O6MeG, but in certain cases their activity on damaged substrates exceeded that of the undamaged controls. Examples of this were O4MeT and 8oxoA-containing substrates being preferred over native bases for AID, and 8oxoG and 8oxoA-containing motifs being preferred over undamaged motifs for A3A and A3B. When damaged nucleotides such as 8oxoG were in the -7, +7, or both -7 & +7 positions a new trend emerged; A3A failed to demonstrate an activity increase relative to undamaged and -1 8oxoG controls while A3B exhibited significantly increased activity for -7, +7, and -7/+7 substrates compared to undamaged and -1 8oxoG controls. AID demonstrated only a mildly significant activity increase for -7 and +7 substrates. As A3B possesses a regulatory domain that A3A and AID lack, experiments with truncated and chimeric versions of A3B and A3A were designed to test the theory that A3 regulatory domains might play a critical role in binding distally damaged ssDNA substrates. While troubleshooting the synthesis of these truncated and chimeric enzymes is ongoing, it is my hope that this domain-based project can continue in the future to assess enzymatic activity on undamaged *vs.* -1 8oxoG *vs.* -7 8oxoG *vs.* +7 8oxoG *vs.* -7/+7 8oxoG-containing substrates, in addition to other substrates containing 8oxoA, O4MeT, 1MeA, and O6MeG.

Troubleshooting attempts were targeted at various stages of the expression protocol. To increase the product yield, I first completed several temperature gradient PCR trials ranging from 48°C-72°C to optimize experimental conditions. Secondly, I attempted to lengthen the extension time; 1 minute, 1 minute 30 seconds, 1 minute 45 seconds, and 2 minutes were assessed. For the same reason, I then attempted increasing the number of cycles (30, 35, and 40), followed by altering the length of time at the annealing temperature for each cycle (30, 45, and 60 seconds, respectively). The 3' EcoRV cut site was later changed to a second BamHI cut site so that insertion into a BamHI/BamHI GST-containing pcDNA 3.1 (+) vector after our research group found this newly generated vector to be more successful in protein expression protocols. In a final effort, trials of both GC-rich and high-fidelity buffers, in addition to high-fidelity PCR, were attempted for BNAC. While these troubleshooting attempts permitted the PCR of Ntd-A3B and Ctd-A3B, they did not aid in the amplification of BNAC. As a result, BNAC was ordered in a vector rather than as a gene fragment; our research group had previously ordered all of our products in a vector, but were trialing a new and supposedly more cost-effective service from GeneBlocks (IDT) at the time. When Ctd-A3B presented with challenges in the ligation and transformation phases, excising the Ctd-A3B out of non-truncated GST-A3B in pcDNA 3.1 (+) was attempted; when the challenges persisted, Ctd-A3B was also ordered in a vector rather than as a gene fragment.

With PCR of Ntd-A3B, Ctd-A3B, and BNAC now possible, further troubleshooting became required at the ligation stage for Ctd-A3B and BNAC. Firstly, I increased the ligation time from the recommended 2 hours to 16 hours, and even attempted a 24-hour ligation for Ctd-A3B when other efforts failed. I trialed a variety of DNA:vector ratios,

including 3:1, 5:1, 8:1, and 10:1. I also attempted each of the aforementioned conditions with and without heat inactivation (incubation at 65°C for 20 minutes) of the restriction enzymes and T4 DNA ligase, and utilized BamHI and T4 DNA ligase from another research laboratory in our department as a comparison to assess the activity of our own enzyme stock (many thanks to Dr. Ken Hirasawa for his support in this endeavour). Additionally, I evaluated if ligation into a PGEX vector followed by a restriction digest and ligation into pcDNA 3.1 (+) would produce viable colonies for Ctd-A3B and BNAC after transformation, but this method appeared ineffective as it did not increase the concentration of pcDNA 3.1 (+) containing the DNA fragment.

In regard to troubleshooting the transformation step, I tried changing the volume of ligation reaction added to the competent cells (1, 5, and 10 µl, respectively). All these volumes were trialed with both XL1-Blue and Top10 competent cells. Eventually colonies containing Ntd-A3B and BNAC grew on LB/ampicillin plates, but as complete sequencing of BNAC could not be completed its presence and orientation in the vector could not be confirmed post-transformation. Ctd-A3B did not produce colonies, even after the additional attempt of allowing the LB/ampicillin plates to remain at 37°C for 24 and 48 hours.

A possible explanation for these difficulties is that the XL1-Blue competent cells mutated the Ctd-A3B and BNAC genomic sequence in such a way that they became toxic to the cells and therefore colonies containing this cytotoxic sequence were negatively selected. This theory is supported by the fact that I experienced no issues in growing colonies containing Ntd-A3B, which is catalytically inactive. Ergo, even if the Ntd-A3B sequence were to be mutated it would be unlikely to experience any changes in its already

non-existent activity. However, as the XL1-Blue strain is known to have a mutagenic rate of 0.005 substitutions per genome per generation (¹³⁷), this possibility seems somewhat statistically unlikely. An avenue to explore the reason behind the problems with expressing the Ctd-A3B and BNAC sequences is *in silico* modelling with a program such as PyMol. Firstly, *in silico* modelling with randomly generated mutations in the intronic region may elucidate if a particular mutation renders the intron strategy ineffective by permitting enzymatic activity in the competent cells. Secondly, this modelling process could also potentially predict if the Ctd-A3B was catalytically active enough to be cytotoxic. If this proved to be the case, and the intron strategy was proven to be sound, Ctd-A3B could potentially be re-ordered with an additional intron.

Should the expression of Ctd-A3B and BNAC become possible, the planned double domain experiments can begin. As predicted earlier in this manuscript, if substrate was undamaged I would expect wild-type A3A to be more active than wild-type A3B as per the evidence from previous damaged base trials (Figures 3, 4, 7, 14, and 15). As Ntd-A3B only plays a regulatory role, and would be expected to exhibit limited or no catalytic activity (²⁵). If A3B relies on its Ntd to increase its overall activity, Ctd-A3B alone should be expected to be less active than full (wild-type) double-domain A3B. Based on previous chimera research (^{4,5,26,28-32}), whichever enzyme is in the Ctd defines overall activity. Therefore, an A3B chimera containing A3A as its Ctd should behave with similar activity to wild-type A3A, but activity should be equal or increased beyond that of wild-type A3A due to the presence of a regulatory domain that A3A normally. However, an alternate possibility is that the added bulk of an additional domain to A3A may actually impede its

substrate binding and lessen its activity, as A3A is already more active than A3B and may not benefit from a regulatory domain.

The structural relationship of the Ntd to enzyme activity may also depend on mutational load. The activity of AID, A3A, and A3B on substrates containing one *versus* two damaged nucleotides was compared by way of analyzing their activity when exposed to -7 8oxoG-containing substrates, +7 8oxoG-containing substrates *versus* -7+7 8oxoG-containing substrates, or -1 8oxoG-containing substrates *versus* -7+7 8oxoG-containing substrates. AID (single domain) and A3A (single domain) were found to exhibit lower activity on -7+7 8oxoG substrates than on +7 8oxoG substrates. AID and A3A also exhibited less activity on -7+7 8oxoG substrates than on -7 8oxoG substrates. AID was less active on -1 8oxoG substrates compared to -7+7 8oxoG substrates, but A3A was more active on -1 8oxoG substrates compared to -7+7 8oxoG substrates. However, A3B (double domain) exhibited the highest activity on -7 8oxoG substrates, followed by -7+7, then +7, and then -1 8oxoG substrates. It therefore appears that single domain APOBECs such as AID and A3A demonstrate less activity overall when exposed to a higher mutational load (i.e. load dependent). On the other hand, A3B activity increased somewhat with mutational load but was still more active when a single damaged nucleotide was present upstream (i.e. somewhat load dependent but may be more location dependent than single domain APOBECs). This analysis supports my hypothesis that the regulatory domain of A3B plays more of a role in binding distally damaged bases. However, it may also indicate that, hypothetically, A3B might trigger repair mechanisms less effectively than AID or A3A; its relative lower activity on substrates containing a greater number of damaged nucleotides may indicate that A3B is less sensitive at

detecting mutational load. Therefore, at this time it is worth speculating that single domain family members such as AID and A3A may play a greater role in the repair whereas double domain family members may play a greater role in carcinogenesis; perhaps my balance hypothesis of repair:carcinogenesis does not apply to each individual APOBEC enzyme, but the enzyme family as a whole.

Although we have not examined genome-wide targeting, our data suggests that, if AID, A3A, or A3B encounter damaged base motifs in one of these ssDNA forms, they would have the catalytic capacity to deaminate some of these substrates even more efficiently than undamaged DNA. Under normal conditions, natural oxidization/alkylation events in the genome are repaired with high efficiency, decreasing the likelihood that AID/A3A/A3B will encounter these occurrences. However, in cancer cells, blatant DNA damage repair deficiencies are present (^{82,138-140}). For instance, 8oxoG, the most prevalent form of oxidized DNA damage in the genome, often causes G:A mismatches within a cellular genome due to its pyrimidine-like behavior (^{91,141}). This accounts for our observation of its favourability for A3A/B as a -1 position base where a pyrimidine is highly favored. In healthy human cells, 8oxoG:A is repaired by OGG1, a glycosylase/AP lyase that excises 8oxoG from the DNA duplex, or by MYH which removes dA from the 8oxoG:A pair; however, these repair pathways can be perturbed as part of the disease processes of cancer (^{91,97}). *O*6-methylguanine (O6MeG), which often results from environmental alkylating DNA damage, is repaired by *O*6-alkylguanine-alkyltransferase (MGMT); the high levels of O6MeG mutations are found within several different types of cancer cells indicate that the MGMT pathway is no longer functioning in these cells (^{90,96}). DNA oxidative damage such as 8oxoG and 8oxoA often lead to the stalling of

transcriptional elongation, potentially exposing oxidized ssDNA regions known to be targets of AID/A3 enzymes (^{142,143}). Oxidation and alkylation of DNA can also induce blockage of replicative polymerases (^{144,145}), thereby affecting the formation of new cells.

On top of the lack of repair mechanisms in cancer cells, the most widely prescribed classes of chemotherapeutics are sources of overburdening DNA alkylation and oxidation, and cause the conversion of native DNA bases into damaged forms such as 8oxoG and O6MeG (^{31-33,41,53-56}). Aside from chemotherapeutics, environmental agents that much of the population is exposed to, such as NNK from cigarette smoke, UV radiation from the sun and tanning salons, dietary nitrosamines from processed meat, and acetaldehyde from alcohol (⁴⁸⁻⁵²), can trigger the formation of oxidized, alkylated, or bulky adduct-containing nucleotide bases.

In consideration of our in vitro data, the possibility emerges that accumulation of damaged nucleotide bases may act through several mechanisms to shift the endogenous patterns of deamination by AID/A3 enzymes. We suggest that this possibility merits further investigation; the idea that AID/A3 enzymes may stimulate repair mechanisms by induction of genomic damage might indicate that the AID/A3 cancer induction pathway may in fact be a balanced continuum with cell repair. This insight may provide a potential target for novel oncological therapies regulating the concentration and activity levels of AID/A3 rather than inhibiting the enzymes entirely.

Since AID/A3 enzymes may be hyperactive in DNA regions containing damaged DNA, it is also tempting to speculate that they may play a role in DNA repair, perhaps through instigating recruitment of mismatch repair or base excision pathways which are known to follow their deaminating activities (¹⁴⁶⁻¹⁴⁸). A3A is directly involved in

activating the DNA damage response in several cell types and regulating key checkpoint proteins in the cell cycle, such as checkpoint kinase 2 and replication protein A (RPA), via phosphorylation; these two proteins signal DNA damage response. Additionally, when A3A is upregulated it may promote cell cycle arrest by targeting the lagging strand of nuclear DNA undergoing replication (¹⁴⁹). A3G has been shown to increase efficiency of DSB repair in Lymphoma cells exposed to ionizing radiation. A3G multimers were also shown to associate with the ssDNA regions and end termini of resected double-stranded breaks (DSBs) (^{106,150}). As a single domain family member (A3A) and a double-domain family member (A3G) have been accounted for in previous repair studies, it is not unreasonable to speculate that other members of the AID/A3 family may play similar roles to maintain the DSB repair rate.

Taken together with our results that AID, A3A and A3B act efficiently on the most frequently found forms of base damage, a potential role for these enzymes in repair also merits further investigation. If AID/A3s are presumed involved in the balance of DNA damage and repair, they could provide a potential chemotherapeutic drug target, possibly one with a specific target rather than systemic effects. Any new knowledge pertaining to how their domain structure interacts with damaged bases in the genome could assist in pharmaceutical development and merits further investigation.

My completed work has several limitations that impact applicability, though the expansion of future trials to include cell-based models could provide solutions (see Chapter 5- Proposal for Future Work). As my studies up to this point involve only *in vitro* work, it cannot accurately predict the results of a mouse *in vivo* trial. However, it could provide valuable insight for the ethical development of these models in the future. The

assessment of AID/A3 activity, which theoretically may act as a beacon for repair enzymes, also does not assess DNA repair directly. Separate trials will need to be designed to assess the impact of increased AID/A3 activity on DNA repair pathways, though the cellular assays may be similar to the proposed uracil assay in some ways. In fact, examining the interaction of AID/A3 with DNA repair enzymes in damaged nuclear environments may prove to be an interesting future direction for our lab and the field of cancer immunology in general.

Chapter 5 - Proposal for Future Work

Summary of Proposed Future Work

The approach to future research following up on the findings of my thesis can be divided into five goals. Firstly, the activity trials of A3A and A3B on undamaged DNA substrates and substrates containing an 8oxoG in the -1, -7, +7, and both -7 and +7 positions relative to the dC (distance damage trials) will be repeated in the laboratory's new location in British Columbia to act as controls for the double domain project. This would also be the optimum time to explore how the activity of A3A and A3B is affected by gradually increasing the distance between the dC and the damaged base in the substrate. For example, substrates with an 8oxoG 2, 3, 4, 5, 6, 8, 9, 10, 14, and 21 nucleotides up- and/or downstream of the dC could be tested. Secondly, these same original substrates (undamaged, -1, -7, +7, and -7/+7) along with any other notable distally damaged substrates from the aforementioned trials will be used to compare the activity of A3A, Ntd-A3B, Ctd-A3B, and a chimera of Ntd-A3B and A3A (BNAC). Thirdly, these activity trials will be expanded to include A3F and A3G (both double-domain enzymes) while investigating the activity of A3A, A3B, Ntd-A3B, Ctd-A3B, BNAC, A3F, and A3G on ssDNA substrates containing damaged bases other than 8oxoG (ex. 8-Oxo-2'-deoxyadenosine, O4-methyl-deoxythymidine, O6-methyl-deoxyguanosine, and 1-methyl-deoxyadenosine). Fourthly, based on the results of Phase III, the synthesis of separate Ntd and Ctd for A3F and A3G will be attempted in order to expand the scope of previous activity trials. Finally, an *in vitro* model would be established to create environmental damage in cell lines. The system will involve HEK-293T cells, with some

lines expressing AID/A3s and some lines having inactive enzymes as a negative control. Environmental mutagens could then be directly introduced into the serum media. After the desired exposure time was achieved, genomic DNA would be extracted from the cell. AID/A3 activity could then be measured via free uridine concentration after uracil DNA-glycosylase (UDG) and an aldehyde-reactive probe (ARP) linked to a fluorescent Cy5-streptavidin tag were added to the DNA. This model could prove beneficial for our lab and others in the field to directly assess APOBEC activity after environmental induction of DNA lesions. I have designed a specific protocol for these procedures using a literature review of similar procedural methods in the field.

Phase I – *Repetition of the Distance Damage Base Experiments and Exploring A3A/B Activity When Exposed to a Range of 8oxoG Locations Within the Substrate*

This stage of experimental work will proceed as per the Materials and Methods section of **Chapter 2**. Undamaged ssDNA and -1 8oxoG (proximal damage) will act as controls for -7, +7, and -7+7 experimental substrates. A3A and A3B could be exposed to substrates with an 8oxoG 2, 3, 4, 5, 6, 8, 9, 10, 14, and 21 nucleotides up- and/or downstream of the dC to further assess how patterns of enzymatic activity, and by extension residue binding (via *in silico* modelling), are affected by the location of a damaged nucleotide. In addition to 8oxoG, other damaged bases such as 8oxoA could be tested. Furthermore, a higher density of damaged bases per substrate (i.e. > 2/oligonucleotide) should be tested, and other substrates shapes besides bubbles (e.g. stem loops, forks, etc.) could be examined. Other than enzyme activity being assessed using the alkaline cleavage assay, enzyme:substrate binding using EMSA and fine analyses of

substrate secondary structure using native gel electrophoresis and nuclease foot-printing will further reveal whether the impact of damaged bases lies in modulating the binding of AID/A3 enzymes to the damaged DNA through direct amino acid:nucleotide interactions between the enzyme and substrate, or whether this effect is a result of the damaged bases altering the overall secondary shape of the DNA in which they reside.

Phase II – Double Domain Damage Experiments with A3A and A3B

Once synthesis and purification of Ntd-A3B (truncated, negative control), Ctd-A3B (truncated), and BNAC (chimeric) has been successful, the purpose of this experiment will be to determine if the regulatory Ntd of soluble domain A3s is responsible for its increased activity on distally damaged substrates (-7, +7, and -7+7). The above truncated and chimeric enzymes will be tested simultaneously with wild-type A3A (single domain) and wild-type A3B (double domain) as controls. Undamaged and -1 8oxoG (proximal damage) will also act as controls for all enzymes (wild-type, truncated, and chimeric). These chimeric enzymes can be further tested on the breadth of new substrates mentioned in Phase I, and based on the results, new chimeric or mutant enzymes can be generated to probe the enzyme structural aspects of this novel phenomenon of increased AID/A3 activity on pre-damaged DNA forms.

Phase III – *Determination of Truncated and Chimeric A3A/A3B Activity on a Variety of Proximally Damaged Substrates*

This stage of experimental work will expand upon previous Phases, as they have only explored Ntd-A3B, Ctd-A3B, and BNAC activity on one proximally damaged substrate (-1 8oxoG) and substrates containing 8oxoG in general. One must wonder if the effect of the Ntd regulatory domain possessed by double-domain A3s will also increase enzyme catalytic activity on substrates containing damaged nucleotide bases such as 8oxoA, O6MeG, and O4MeT. To examine this, wild type A3A, wild type A3B, Ntd-A3B, Ctd-A3B, and BNAC will be exposed to undamaged ssDNA substrate and ssDNA substrates containing 8oxoG, 8oxoA, O6MeG, 1MeA, and O4MeT in the -1 position as controls. The -1 8oxoG could act as a positive control, as the catalytic efficiency for it will already be known from Phase II of experiments. The other -1 damaged substrates will also act as controls for substrates containing 8oxoG, 8oxoA, O6MeG, and O4MeT in the -7, +7, and -7+7 positions relative to the dC. As 8oxoA and O6MeG (in addition to our control 8oxoG) were preferred substrates of A3A and A3B in proximally damaged base trials, and adenine and guanine share a similar dual-ring structure, I expect the Ctd regulatory domain to play a role in increased binding. Ergo, BNAC should exhibit increased activity on 8oxoA and O6MeG -7, +7, and -7+7 substrates compared to wild-type A3A, and Ctd-A3B should exhibit less activity on these substrates relative to wild-type A3B. 1MeA and O4MeT did not significantly increase A3A or A3B activity in proximally damaged base trials, so I hypothesize they will be unlikely to do so in a distally damaged location as well.

Phase IV – *Synthesis and Analysis of Truncated A3F/A3G Activity on Proximally and Distally Damaged Substrates*

Once A3A and A3B activity has been thoroughly explored, determining if other double-domain A3s follow the same pattern of their regulatory domain increasing activity on distally damaged substrates will strengthen the theory. Therefore, Ntd-A3F, Ctd-A3F, Ntd-A3G, and Ctd-A3G will be synthesized and used alongside wild-type A3F (double-domain) and wild-type A3G (double-domain) as controls. For substrates, undamaged and -1 8oxoG, 8oxoA, O6MeG, 1MeA and O4MeT would act as controls for the distally damaged substrates (-7, +7, and -7+7 8oxoG, 8oxoA, O6MeG, 1MeA, and O4MeT substrates). As before in Phase III, each -1 damaged substrate will act as a counterpart control for its distally damaged substrate. Assuming that the presence of a regulatory domain is key for A3 activity on distally damaged ssDNA substrates, previous trends should hold true for A3F and A3G as they are both double-domain enzymes. Ergo, Ntd-A3F and Ntd-A3G (negative controls) will exhibit near zero activity on proximally and distally damaged substrates. Ctd-A3F and Ctd-A3G will display lesser activity than wild-type A3F and wild-type A3G (which contain a regulatory Ntd domain in addition to their catalytic Ctd domain), respectively, on distally damaged substrates compared to proximally damaged or undamaged control substrates. Ctd-A3F and Ctd-A3F should also be active on proximally damaged substrates, but not as active as their wild-type counterparts.

Phase V – Development of More Sophisticated In Vitro and Ex Vivo Models for Induction of DNA Damage by Environmental Factors

Phase V marks the need to advance beyond test tube studies and examine how AID/A3 activity is affected by the presence of damaged DNA bases in a cellular environment where some repair enzymes may still be active. Therefore, a protocol to induce environmental damage in cell lines will be attempted. The damaging agents utilized will be NNK, H₂O₂, acetaldehyde, and UVA/B light, as these are currently readily available in our lab and are representative of damage due to smoking, ROS, alcohol, and tanning/sun exposure, respectively. If this work is successful, placing the cell cultures in radiation oncology machines or exposing them to oncology medications such as cis-platin may prove intriguing. The most direct way to measure AID/A3 activity in cells will be through quantitation of uridine, as this is the product of AID/A3 reactions (see ***Proposed Specific Protocol for Phase V***). HEK 293 T-cell lines expressing the desired active AID/A3 will be used, while cell lines expressing an inactive form of the desired AID/A3 will serve as a control. A further control will include cell cultures (containing both active and inactive AID/A3s, separately) that are not exposed to any environmental mutagen. As repair enzymes should be at least somewhat inhibited in cancer cells (such as HEK 293 T cells), I hypothesize that a greater quantity of uridines will be detectable in the cell cultures that have been exposed to environmental mutagens; this directly translates to the fact that AID/A3s have been demonstrated to be more active on damaged substrates. As a mid-way point between the studies described here and those proposed in Phase V, one could attempt *in vitro* studies on larger DNA sequences such as plasmids to establish the methodology for quantification of base damage as well as the interplay between uracil

generation by AID/A3s and pre-existing base damage. In these assays, DNA damage by oxidation or alkylation would be induced in plasmid DNA, which would then be incubated with AID/A3 enzymes, followed by measurement of AID/A3 activity via quantification of uracil levels.

Proposed Specific Protocol for Phase V

Proposed Specific Protocol - *Induction of DNA Damage in Cells with Environmental Mutagens*

HEK 293 T cells will be cultured in modified Eagle's medium at 37°C in a 5% CO₂ atmosphere. Cells will be seeded into 75 cm² flasks of complete medium and permitted to grow for 24 hours. These cells would then be exposed to varying levels of environmental mutagens as desired for an additional 24 hours before having their media removed and being washed with 1x PBS. Further culture time could be carried out for the study of DNA repair if desired before cells are detached using trypsin-Ethylenediaminetetraacetic acid and harvested by centrifugation. Planned mutagen exposure concentrations vary depending on the substance. NNK at concentrations of 5, 10, or 25 µM ⁽¹⁵¹⁾ might prove suitable. H₂O₂ (Sigma-Aldrich, 30% (w/w) in H₂O) will be diluted with PBS to yield final concentrations of 2000 µM, 1500 µM, and 100 µM ⁽¹⁵²⁾. 0.01-40 mM ranges of acetaldehyde could be trialed with 25 mg of DNA in 1.5 mL of 0.1M phosphate buffer pH 7 ^(48,49). Cells in PBS could be exposed to 6 J cm⁻² (80 mW cm⁻²) and 60 mJ cm⁻² (1.44 mW cm⁻²) UVA, and 280–370 nm UVB radiation ⁽¹⁵³⁻¹⁵⁵⁾.

Proposed Specific Protocol - *Design of a Uridine Assay: Isolation and Preparation of Cellular DNA*

The suggested protocol will be based on that used in other labs (¹⁵⁶⁻¹⁵⁸), though may require optimization for use in our lab. HEK 293 T cells will be suspended (with AID/A3s active or inactive as desired) in RPMI 1640 HyClone media supplemented with 10% fetal bovine serum (HyClone) and 1% penicillin-streptomycin. Induction of damage with environmental mutagens including appropriate negative controls, if desired, would occur next. Genomic DNA would subsequently be extracted with DNAzol Reagent, then incubated with 100 mM methoxyamine (a blocking agent) in 50 mM Tris-HCl at 37°C for 2 hours. DNA would then be precipitated with 4 volumes of 100% ice-cold ethanol and 7.5% volume of 4 M NaCl to remove the methoxyamine. DNA could be resuspended in TE pH 7.6 if required (recommended for UV exposure trials). Next, the DNA will be treated with *E. coli* 0.2 units uracil DNA-glycosylase (UDG) followed by a 2 mM aldehyde-reactive probe (ARP), at 37°C for 15 minutes each. UDG removal may require extraction with phenol-chloroform before ethanol precipitation, storage in TE buffer pH 7.6, and removal of artifactual damage using G-50 or Qiagen gravity tip columns (¹⁵⁶⁻¹⁵⁸).

Proposed Specific Protocol - *Design of a Uridine Assay: Membrane Hybridization*

Nitrocellulose membranes will be pre-equilibrated with StartingBlock buffer and washed with 500 uL of ammonium acetate. DNA samples and standards (400 uL, 100 ng) will be spotted onto positively charged nylon membranes using a vacuum-filtration apparatus such as the Gibco Filtration Manifold system. Membranes will then be

washed in 5x saline sodium citrate (SSC) for 15 minutes at 37°C and baked under vacuum at 80°C for 30 minutes. Baked membranes will subsequently be incubated in 40 mL of hybridization buffer (20 mM Tris, pH 7.5; 0.1 M NaCl; 1 mM Ethylenediaminetetraacetic acid; 0.5% casein w/v; 0.25% BSA w/v; 0.1% Tween-20 v/v) for 30 minutes at room temperature. The membrane will then undergo a second incubation in a solution of 100 uL Cy5-streptavidin (a chemiluminescent reagent) made up to 40 mL with fresh hybridization buffer at room temperature for 45 minutes. The membrane will then be washed 3x for 5 minutes each with TBS-T pH 7.5 (25 mM Tris or 20 mM, 3 mM KCl, 140 mM NaCl, 1% Tween 20). Incubation in enhanced chemiluminescent (ECL) reagent (Pierce, Rockford, IL or ThermoFisher Scientific) at room temperature for 5 minutes will then permit scanning of the moist membrane using a phosphorimager such as a Typhoon 9210. The fluorescent signal will be directed upwards, and a 50 μ M setting will be used to achieve highest resolution. Exposure time will be 4 minutes. Finally, DNA can be quantified from each sample using SYBR gold dye and a Tecan Genios microplate reader. Analysis of Cy5 fluorescence will be carried out using ImageJ or Quantity One software. A standard curve of Cy5 fluorescence *versus* uracil amounts will be generated, with raw fluorescence numbers adjusted for background fluorescence signal. Uracil concentration of the unknowns will then be interpolated from the standard curve, and normalized by dividing each unknown value by the “control” DNA value to calculate the number of uracils per 10⁶ base pairs (bp) (1 million bp = $\sim 1.05 \times 10^{-15}$ g) (¹⁵⁶⁻¹⁵⁸).

Proposed Specific Protocol - *Design of a Uridine Assay: Preparation of Uracil Standards*

A 10 pmol oligonucleotide duplex such as 5'-T37UT37/5'-A37GA37 or U-containing substrates our lab already possesses will be used as uracil standards, with suggested concentrations being 4 µg, 2 µg, 1 µg, 0.5 µg, 0.25 µg, 0.125 µg, and 0 µg. After blocking with methoxyamine, appropriate controls can be established by exposing half of each dilution to UDG and not introducing UDG into the other half. The standards will be probed with ARP and immobilized onto a nitrocellulose membrane. Band intensity will be quantified using an ECL/ChemiImagerTM system, and utilized to generate a standard curve (¹⁵⁶⁻¹⁵⁸).

Bibliography

1. Seifert M, Kuppers R. Human memory B cells. *Leukemia* 2016;30:2283-92.
2. Rajewsky K. Clonal selection and learning in the antibody system. *Nature* 1996;381:751-8.
3. Hamilton CE, Papavasiliou FN, Rosenberg BR. Diverse functions for DNA and RNA editing in the immune system. *RNA biology* 2010;7:220-8.
4. Ito F, Fu Y, Kao SA, Yang H, Chen XS. Family-Wide Comparative Analysis of Cytidine and Methylcytosine Deamination by Eleven Human APOBEC Proteins. *Journal of molecular biology* 2017;429:1787-99.
5. Bransteitter R, Prochnow C, Chen XS. The current structural and functional understanding of APOBEC deaminases. *Cellular and molecular life sciences : CMLS* 2009;66:3137-47.
6. Muramatsu M KK, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced deoxycytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 2000;102:553-63.
7. Martin A, Scharff MD. AID and mismatch repair in antibody diversification. *Nature reviews Immunology* 2002;2:605-14.
8. Chaudhuri J, Basu U, Zarrin A, et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Advances in immunology* 2007;94:157-214.
9. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annual review of immunology* 2008;26:261-92.
10. Navaratnam N, Morrison JR, Bhattacharya S, et al. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *The Journal of biological chemistry* 1993;268:20709-12.
11. Prochnow C, Bransteitter R, Klein MG, Goodman MF, Chen XS. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* 2007;445:447-51.
12. Lada AG, Krick CF, Kozmin SG, et al. Mutator effects and mutation signatures of editing deaminases produced in bacteria and yeast. *Biochemistry Biokhimiia* 2011;76:131-46.
13. Harris RS, Petersen-Mahrt SK, Neuberger MS. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Molecular cell* 2002;10:1247-53.
14. Grant M, Larijani M. Evasion of adaptive immunity by HIV through the action of host APOBEC3G/F enzymes. *AIDS research and therapy* 2017;14:44.
15. Conticello SG, Langlois MA, Yang Z, Neuberger MS. DNA deamination in immunity: AID in the context of its APOBEC relatives. *Advances in immunology* 2007;94:37-73.
16. Larijani M, Martin A. The biochemistry of activation-induced deaminase and its physiological functions. *Seminars in immunology* 2012;24:255-63.
17. Burns MB LL, Carpenter MA, et al. . APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 2013;494:366-70.

18. Roberts SA LM, Klimczak LJ, et al. . An APOBEC deoxycytidine deaminase mutagenesis pattern is widespread in human cancers. *Nature genetics* 2013;45:970-6.
19. Sawai Y KY, Shimizu T, et al. . Activation-Induced Deoxycytidine deaminase Contributes to Pancreatic Tumorigenesis by Inducing Tumor-Related Gene Mutations. *Cancer research* 2015;75:3292-301.
20. Starrett GJ, Luengas EM, McCann JL, et al. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nature communications* 2016;7:12918.
21. Suzuki T, Kamiya H. Mutations induced by 8-hydroxyguanine (8-oxo-7,8-dihydroguanine), a representative oxidized base, in mammalian cells. *Genes and environment : the official journal of the Japanese Environmental Mutagen Society* 2017;39:2.
22. Rogozin IB, Roche-Lima A, Lada AG, et al. Nucleotide Weight Matrices Reveal Ubiquitous Mutational Footprints of AID/APOBEC Deaminases in Human Cancer Genomes. *Cancers* 2019;11.
23. Klemm L, Duy C, Iacobucci I, et al. The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. *Cancer cell* 2009;16:232-45.
24. Shinmura K, Igarashi H, Goto M, et al. Aberrant expression and mutation-inducing activity of AID in human lung cancer. *Annals of surgical oncology* 2011;18:2084-92.
25. Matsumoto Y MH, Kinoshita K, Niwa Y, Sakai Y, Chiba T. Upregulation of activation-induced deoxycytidine deaminase causes genetic aberrations at the CDKN2b-CDKN2a in gastric cancer. *Gastroenterology* 2010;139:1984-94.
26. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415-21.
27. Roberts SA, Gordenin DA. Hypermutation in human cancer genomes: footprints and mechanisms. *Nature reviews Cancer* 2014;14:786-800.
28. Paz-Elizur T, Krupsky M, Blumenstein S, Elinger D, Schechtman E, Livneh Z. DNA repair activity for oxidative damage and risk of lung cancer. *Journal of the National Cancer Institute* 2003;95:1312-9.
29. Cooke MS, Evans MD, Dizdaroglu M, Lunec J. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2003;17:1195-214.
30. Jiang Y, Liu Y, Hu H. Studies on DNA Damage Repair and Precision Radiotherapy for Breast Cancer. *Advances in experimental medicine and biology* 2017;1026:105-23.
31. Cooke MS, Evans MD. 8-Oxo-deoxyguanosine: reduce, reuse, recycle? *Proceedings of the National Academy of Sciences of the United States of America* 2007;104:13535-6.
32. Fortini P, Pascucci B, Parlanti E, D'Errico M, Simonelli V, Dogliotti E. 8-Oxoguanine DNA damage: at the crossroad of alternative repair pathways. *Mutation research* 2003;531:127-39.
33. Dizdaroglu M. Oxidatively induced DNA damage: mechanisms, repair and disease. *Cancer letters* 2012;327:26-47.

34. Talhaoui I, Couve S, Ishchenko AA, Kunz C, Schar P, Saparbaev M. 7,8-Dihydro-8-oxoadenine, a highly mutagenic adduct, is repaired by *Escherichia coli* and human mismatch-specific uracil/thymine-DNA glycosylases. *Nucleic acids research* 2013;41:912-23.
35. Kasai H. Analysis of a form of oxidative DNA damage, 8-hydroxy-2'-deoxyguanosine, as a marker of cellular oxidative stress during carcinogenesis. *Mutation research* 1997;387:147-63.
36. Viel A, Bruselles A, Meccia E, et al. A Specific Mutational Signature Associated with DNA 8-Oxoguanine Persistence in MUTYH-defective Colorectal Cancer. *EBioMedicine* 2017;20:39-49.
37. Halliwell B. Can oxidative DNA damage be used as a biomarker of cancer risk in humans? Problems, resolutions and preliminary results from nutritional supplementation studies. *Free radical research* 1998;29:469-86.
38. Rai P, Onder TT, Young JJ, et al. Continuous elimination of oxidized nucleotides is necessary to prevent rapid onset of cellular senescence. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106:169-74.
39. Abbott PJ, Saffhill R. DNA-synthesis with methylated poly(dA-dT) templates: possible role of O4-methylthymine as a pro-mutagenic base. *Nucleic acids research* 1977;4:761-9.
40. Kaina B, Fritz G, Mitra S, Coquerelle T. Transfection and expression of human O6-methylguanine-DNA methyltransferase (MGMT) cDNA in Chinese hamster cells: the role of MGMT in protection against the genotoxic effects of alkylating agents. *Carcinogenesis* 1991;12:1857-67.
41. Noonan EM SD, Yaffe MB, Lauffenburger DA, Samson LD. O6-Methylguanine DNA lesions induce an intra-S-Phase arrest from which cells exit into apoptosis governed by early and late multi-pathway signaling network activation. *Integrative biology : quantitative biosciences from nano to macro* 2012;4:1237-55.
42. Pauly GT, Hughes SH, Moschel RC. Comparison of mutagenesis by O6-methyl- and O6-ethylguanine and O4-methylthymine in *Escherichia coli* using double-stranded and gapped plasmids. *Carcinogenesis* 1998;19:457-61.
43. Pourquier P, Waltman JL, Urasaki Y, et al. Topoisomerase I-mediated cytotoxicity of N-methyl-N'-nitro-N-nitrosoguanidine: trapping of topoisomerase I by the O6-methylguanine. *Cancer research* 2001;61:53-8.
44. Altshuler KB, Hodes CS, Essigmann JM. Intrachromosomal probes for mutagenesis by alkylated DNA bases replicated in mammalian cells: a comparison of the mutagenicities of O4-methylthymine and O6-methylguanine in cells with different DNA repair backgrounds. *Chemical research in toxicology* 1996;9:980-7.
45. Larson K, Sahm J, Shenkar R, Strauss B. Methylation-induced blocks to in vitro DNA replication. *Mutation research* 1985;150:77-84.
46. Delaney JC, Essigmann JM. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in *alkB* *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101:14051-6.
47. Yang H, Lam SL. Effect of 1-methyladenine on thermodynamic stabilities of double-helical DNA structures. *FEBS letters* 2009;583:1548-53.

48. Wang M, McIntee EJ, Cheng G, Shi Y, Villalta PW, Hecht SS. Identification of DNA adducts of acetaldehyde. *Chemical research in toxicology* 2000;13:1149-57.
49. Garaycoechea JI, Crossan GP, Langevin F, et al. Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature* 2018;553:171-7.
50. Lieberman HB. DNA damage repair and response proteins as targets for cancer therapy. *Current medicinal chemistry* 2008;15:360-7.
51. Lijinsky W. N-Nitroso compounds in the diet. *Mutation research* 1999;443:129-38.
52. Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* 2002;21:7435-51.
53. Kondo N, Takahashi A, Ono K, Ohnishi T. DNA damage induced by alkylating agents and repair pathways. *Journal of nucleic acids* 2010;2010:543531.
54. Cheung-Ong K, Giaever G, Nislow C. DNA-damaging agents in cancer chemotherapy: serendipity and chemical biology. *Chemistry & biology* 2013;20:648-59.
55. Margison GP, Santibanez Koref MF, Povey AC. Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine. *Mutagenesis* 2002;17:483-7.
56. Ding Y, Wang H, Niu J, et al. Induction of ROS overload by alantolactone prompts oxidative DNA damage and apoptosis in colorectal cancer cells. *International journal of molecular sciences* 2016;17:558.
57. Torgovnick A, Schumacher B. DNA repair mechanisms in cancer development and therapy. *Frontiers in genetics* 2015;6:157.
58. Barnes DE, Lindahl T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annual review of genetics* 2004;38:445-76.
59. Fu D, Calvo JA, Samson LD. Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nature reviews Cancer* 2012;12:104-20.
60. Lord CJ, Ashworth A. The DNA damage response and cancer therapy. *Nature* 2012;481:287-94.
61. Cook JA, Gius D, Wink DA, Krishna MC, Russo A, Mitchell JB. Oxidative stress, redox, and the tumor microenvironment. *Seminars in radiation oncology* 2004;14:259-66.
62. Miyake H, Hara I, Kamidono S, Eto H. Oxidative DNA damage in patients with prostate cancer and its response to treatment. *The Journal of urology* 2004;171:1533-6.
63. Weiss JM, Goode EL, Ladiges WC, Ulrich CM. Polymorphic variation in hOGG1 and risk of cancer: a review of the functional and epidemiologic literature. *Molecular carcinogenesis* 2005;42:127-41.
64. Diakowska D, Lewandowski A, Kopec W, Diakowski W, Chrzanowska T. Oxidative DNA damage and total antioxidant status in serum of patients with esophageal squamous cell carcinoma. *Hepato-gastroenterology* 2007;54:1701-4.
65. Tanaka H, Fujita N, Sugimoto R, et al. Hepatic oxidative DNA damage is associated with increased risk for hepatocellular carcinoma in chronic hepatitis C. *British journal of cancer* 2008;98:580-6.
66. Kesarwala AH, Krishna MC, Mitchell JB. Oxidative stress in oral diseases. *Oral diseases* 2016;22:9-18.

67. Witz G. Active oxygen species as factors in multistage carcinogenesis. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine* 1991;198:675-82.
68. Balaban RS, Nemoto S, Finkel T. Mitochondria, oxidants, and aging. *Cell* 2005;120:483-95.
69. Zhu X, Su B, Wang X, Smith MA, Perry G. Causes of oxidative stress in Alzheimer disease. *Cellular and molecular life sciences : CMLS* 2007;64:2202-10.
70. Nishikawa T, Araki E. Impact of mitochondrial ROS production in the pathogenesis of diabetes mellitus and its complications. *Antioxidants & redox signaling* 2007;9:343-53.
71. Niwa Y, Miyake S, Sakane T, Shingu M, Yokoyama M. Auto-oxidative damage in Behcet's disease--endothelial cell damage following the elevated oxygen radicals generated by stimulated neutrophils. *Clinical and experimental immunology* 1982;49:247-55.
72. Buldanlioglu S, Turkmen S, Ayabakan HB, et al. Nitric oxide, lipid peroxidation and antioxidant defence system in patients with active or inactive Behcet's disease. *The British journal of dermatology* 2005;153:526-30.
73. Onur E, Kabaroglu C, Inanir I, et al. Oxidative stress impairs endothelial nitric oxide levels in Behcets' disease. *Cutaneous and ocular toxicology* 2011;30:217-20.
74. Pagano G, Castello G, Pallardo FV. Sjogren's syndrome-associated oxidative stress and mitochondrial dysfunction: prospects for chemoprevention trials. *Free radical research* 2013;47:71-3.
75. Norheim KB, Jonsson G, Harboe E, Hanasand M, Goransson L, Omdal R. Oxidative stress, as measured by protein oxidation, is increased in primary Sjogren's syndrome. *Free radical research* 2012;46:141-6.
76. Wakamatsu TH, Dogru M, Matsumoto Y, et al. Evaluation of lipid oxidative stress status in Sjogren syndrome patients. *Investigative ophthalmology & visual science* 2013;54:201-10.
77. Dizdaroglu M. Oxidatively induced DNA damage and its repair in cancer. *Mutation research Reviews in mutation research* 2015;763:212-45.
78. Tubbs A, Nussenzweig A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell* 2017;168:644-56.
79. Mao P, Brown AJ, Malc EP, et al. Genome-wide maps of alkylation damage, repair, and mutagenesis in yeast reveal mechanisms of mutational heterogeneity. *Genome research* 2017;27:1674-84.
80. Bartkova J, Hamerlik P, Stockhausen MT, et al. Replication stress and oxidative damage contribute to aberrant constitutive activation of DNA damage signalling in human gliomas. *Oncogene* 2010;29:5095-102.
81. Bartkova J, Hoei-Hansen CE, Krizova K, et al. Patterns of DNA damage response in intracranial germ cell tumors versus glioblastomas reflect cell of origin rather than brain environment: implications for the anti-tumor barrier concept and treatment. *Molecular oncology* 2014;8:1667-78.
82. Frick A, Khare V, Paul G, et al. Overt Increase of Oxidative Stress and DNA Damage in Murine and Human Colitis and Colitis-Associated Neoplasia. *Molecular cancer research : MCR* 2018;16:634-42.

83. Maciag A, Sithanandam G, Anderson LM. Mutant K-rasV12 increases COX-2, peroxides and DNA damage in lung cells. *Carcinogenesis* 2004;25:2231-7.
84. Romanowska M, Maciag A, Smith AL, et al. DNA damage, superoxide, and mutant K-ras in human lung adenocarcinoma cells. *Free radical biology & medicine* 2007;43:1145-55.
85. de Miranda NF, Peng R, Georgiou K, et al. DNA repair genes are selectively mutated in diffuse large B cell lymphomas. *The Journal of experimental medicine* 2013;210:1729-42.
86. Lahtz C, Pfeifer GP. Epigenetic changes of DNA repair genes in cancer. *Journal of molecular cell biology* 2011;3:51-8.
87. Dietlein F, Thelen L, Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. *Trends in genetics : TIG* 2014;30:326-39.
88. Mambo E, Nyaga SG, Bohr VA, Evans MK. Defective repair of 8-hydroxyguanine in mitochondria of MCF-7 and MDA-MB-468 human breast cancer cell lines. *Cancer research* 2002;62:1349-55.
89. Trzeciak AR, Nyaga SG, Jaruga P, Lohani A, Dizdaroglu M, Evans MK. Cellular repair of oxidatively induced DNA base lesions is defective in prostate cancer cell lines, PC-3 and DU-145. *Carcinogenesis* 2004;25:1359-70.
90. Crosbie PA, McGown G, Thorncroft MR, et al. Association between lung cancer risk and single nucleotide polymorphisms in the first intron and codon 178 of the DNA repair gene, O6-alkylguanine-DNA alkyltransferase. *International journal of cancer* 2008;122:791-5.
91. Delaney S, Jarem DA, Volle CB, Yennie CJ. Chemical and biological consequences of oxidatively damaged guanine in DNA. *Free radical research* 2012;46:420-41.
92. Esposito MT, So CW. DNA damage accumulation and repair defects in acute myeloid leukemia: implications for pathogenesis, disease progression, and chemotherapy resistance. *Chromosoma* 2014;123:545-61.
93. Scott TL, Rangaswamy S, Wicker CA, Izumi T. Repair of oxidative DNA damage and cancer: recent progress in DNA base excision repair. *Antioxidants & redox signaling* 2014;20:708-26.
94. Smith CG, West H, Harris R, et al. Role of the oxidative DNA damage repair gene OGG1 in colorectal tumorigenesis. *Journal of the National Cancer Institute* 2013;105:1249-53.
95. Woods D, Turchi JJ. Chemotherapy induced DNA damage response: convergence of drugs and pathways. *Cancer biology & therapy* 2013;14:379-89.
96. Povey AC, Hall CN, Cooper DP, O'Connor PJ, Margison GP. Determinants of O(6)-alkylguanine-DNA alkyltransferase activity in normal and tumour tissue from human colon and rectum. *International journal of cancer* 2000;85:68-72.
97. Al-Tassan N, Chmiel NH, Maynard J, et al. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nature genetics* 2002;30:227-32.
98. Katuwal M, Wang Y, Schmitt K, et al. Cellular HIV-1 inhibition by truncated old world primate APOBEC3A proteins lacking a complete deaminase domain. *Virology* 2014;468-470:532-44.

99. Sawyer SL, Emerman M, Malik HS. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS biology* 2004;2:E275.
100. Shi K, Carpenter MA, Banerjee S, et al. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nature structural & molecular biology* 2017;24:131-9.
101. Xiao X, Yang H, Arutiunian V, et al. Structural determinants of APOBEC3B non-catalytic domain for molecular assembly and catalytic regulation. *Nucleic acids research* 2017;45:7494-506.
102. King JJ, Larijani M. A Novel Regulator of Activation-Induced Cytidine Deaminase/APOBECs in Immunity and Cancer: Schrodinger's CATalytic Pocket. *Frontiers in immunology* 2017;8:351.
103. Kouno T, Silvas TV, Hilbert BJ, et al. Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nature communications* 2017;8:15024.
104. Abdouni H, King JJ, Suliman M, Quinlan M, Fifield H, Larijani M. Zebrafish AID is capable of deaminating methylated deoxycytidines. *Nucleic acids research* 2013;41:5457-68.
105. Larijani M, Martin A. Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. *Molecular and cellular biology* 2007;27:8038-48.
106. Nowarski R, Prabhu P, Kenig E, Smith Y, Britan-Rosich E, Kotler M. APOBEC3G inhibits HIV-1 RNA elongation by inactivating the viral trans-activation response element. *Journal of molecular biology* 2014;426:2840-53.
107. Li J, Chen Y, Li M, et al. APOBEC3 multimerization correlates with HIV-1 packaging and restriction activity in living cells. *Journal of molecular biology* 2014;426:1296-307.
108. Hache G, Liddament MT, Harris RS. The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *The Journal of biological chemistry* 2005;280:10920-4.
109. Logue EC, Bloch N, Dhuey E, et al. A DNA sequence recognition loop on APOBEC3A controls substrate specificity. *PloS one* 2014;9:e97062.
110. Goila-Gaur R, Khan MA, Miyagi E, Kao S, Strebel K. Targeting APOBEC3A to the viral nucleoprotein complex confers antiviral activity. *Retrovirology* 2007;4:61.
111. Pak V, Heidecker G, Pathak VK, Derse D. The role of amino-terminal sequences in cellular localization and antiviral activity of APOBEC3B. *Journal of virology* 2011;85:8538-47.
112. McDougale RM, Hultquist JF, Stabell AC, Sawyer SL, Harris RS. D316 is critical for the enzymatic activity and HIV-1 restriction potential of human and rhesus APOBEC3B. *Virology* 2013;441:31-9.
113. Hakata Y, Landau NR. Reversed functional organization of mouse and human APOBEC3 cytidine deaminase domains. *The Journal of biological chemistry* 2006;281:36624-31.
114. Borzooee F, Larijani M. *Pichia pastoris* as a host for production and isolation of mutagenic AID/APOBEC enzymes involved in cancer and immunity. *New biotechnology* 2019;51:67-79.

115. Cole CG, McCann OT, Collins JE, et al. Finishing the finished human chromosome 22 sequence. *Genome biology* 2008;9:R78.
116. Akre MK, Starrett GJ, Quist JS, et al. Mutation Processes in 293-Based Clones Overexpressing the DNA Cytosine Deaminase APOBEC3B. *PloS one* 2016;11:e0155391.
117. Land AM, Wang J, Law EK, et al. Degradation of the cancer genomic DNA deaminase APOBEC3B by SIV Vif. *Oncotarget* 2015;6:39969-79.
118. Nott A, Meislin SH, Moore MJ. A quantitative analysis of intron effects on mammalian gene expression. *Rna* 2003;9:607-17.
119. Zhou Z, Luo MJ, Straesser K, Katahira J, Hurt E, Reed R. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* 2000;407:401-5.
120. Hamer DH, Smith KD, Boyer SH, Leder P. SV40 recombinants carrying rabbit beta-globin gene coding sequences. *Cell* 1979;17:725-35.
121. R D Palmiter EPS, M R Avarbock, D D Allen, R L Brinster. Heterologous introns can enhance expression of transgenes in mice. *PNAS* 1991;88:478-82.
122. P Gruss CJL, R Dhar, G Khoury. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. *PNAS* 1979;76:4317-21.
123. Hervé Le Hir AN, Melissa J. Moore. How introns influence and enhance eukaryotic gene expression. *Trends in Biochemical Sciences* 2003;28:215-20.
124. Caval V, Bouzidi MS, Suspene R, et al. Molecular basis of the attenuated phenotype of human APOBEC3B DNA mutator enzyme. *Nucleic acids research* 2015;43:9340-9.
125. Caval V, Suspene R, Shapira M, Vartanian JP, Wain-Hobson S. A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3'UTR enhances chromosomal DNA damage. *Nature communications* 2014;5:5129.
126. Quinlan EM, King JJ, Amemiya CT, Hsu E, Larijani M. Biochemical Regulatory Features of Activation-Induced Cytidine Deaminase Remain Conserved from Lampreys to Humans. *Molecular and cellular biology* 2017;37.
127. Dancyger AM, King JJ, Quinlan MJ, et al. Differences in the enzymatic efficiency of human and bony fish AID are mediated by a single residue in the C terminus modulating single-stranded DNA binding. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2012;26:1517-25.
128. Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. *The Journal of experimental medicine* 2003;197:1291-6.
129. Yu K, Huang FT, Lieber MR. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *The Journal of biological chemistry* 2004;279:6496-500.
130. Carpenter MA, Rajagurubandara E, Wijesinghe P, Bhagwat AS. Determinants of sequence-specificity within human AID and APOBEC3G. *DNA repair* 2010;9:579-87.
131. Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ, Stivers JT. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *The Journal of biological chemistry* 2009;284:22898-904.

132. Aguiar PH, Furtado C, Repoles BM, et al. Oxidative stress and DNA lesions: the role of 8-oxoguanine lesions in Trypanosoma cruzi cell viability. PLoS neglected tropical diseases 2013;7:e2279.
133. Abdouni HS, King JJ, Ghorbani A, Fifield H, Berghuis L, Larijani M. DNA/RNA hybrid substrates modulate the catalytic activity of purified AID. Molecular immunology 2018;93:94-106.
134. King JJ, Manuel CA, Barrett CV, et al. Catalytic pocket inaccessibility of activation-induced cytidine deaminase is a safeguard against excessive mutagenic activity. Structure 2015;23:615-27.
135. Sébastien Landry IN, Daniel C Linfesty, Matthew D Weitzman. APOBEC3A can activate the DNA damage response and cause cell - cycle arrest. EMBO reports 2011;12:444-50.
136. Li X, Caval V, Wain-Hobson S, Vartanian JP. Elephant APOBEC3A cytidine deaminase induces massive double-stranded DNA breaks and apoptosis. Scientific reports 2019;9:728.
137. Badran AH, Liu DR. Development of potent in vivo mutagenesis plasmids with broad mutational spectra. Nature communications 2015;6:8425.
138. Chae YK, Anker JF, Carneiro BA, et al. Genomic landscape of DNA repair genes in cancer. Oncotarget 2016;7:23312-21.
139. Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. Nature reviews Cancer 2012;12:801-17.
140. Jeggo PA, Pearl LH, Carr AM. DNA repair, genome stability and cancer: a historical perspective. Nature reviews Cancer 2016;16:35-42.
141. Shibutani S, Takeshita M, Grollman AP. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. Nature 1991;349:431-4.
142. Canugovi C, Samaranayake M, Bhagwat AS. Transcriptional pausing and stalling causes multiple clustered mutations by human activation-induced deaminase. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 2009;23:34-44.
143. Kuraoka I, Endou M, Yamaguchi Y, Wada T, Handa H, Tanaka K. Effects of endogenous DNA base lesions on transcription elongation by mammalian RNA polymerase II. Implications for transcription-coupled DNA repair and transcriptional mutagenesis. The Journal of biological chemistry 2003;278:7294-9.
144. Minca EC, Kowalski D. Replication fork stalling by bulky DNA damage: localization at active origins and checkpoint modulation. Nucleic acids research 2011;39:2610-23.
145. E. M. Not breathing is not an option: How to deal with oxidative DNA damage. DNA repair 2017;59:82-105.
146. Siriwardena SU, Chen K, Bhagwat AS. Functions and Malfunctions of Mammalian DNA-Cytosine Deaminases. Chemical reviews 2016;116:12688-710.
147. Pilzecker B, Jacobs H. Mutating for Good: DNA Damage Responses During Somatic Hypermutation. Frontiers in immunology 2019;10:438.
148. Methot SP, Di Noia JM. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. Advances in immunology 2017;133:37-87.

149. Landry S, Narvaiza I, Linfesty DC, Weitzman MD. APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. *EMBO reports* 2011;12:444-50.
150. Nowarski R, Kotler M. APOBEC3 cytidine deaminases in double-strand DNA break repair and cancer promotion. *Cancer research* 2013;73:3494-8.
151. Leng J, Wang Y. Liquid Chromatography-Tandem Mass Spectrometry for the Quantification of Tobacco-Specific Nitrosamine-Induced DNA Adducts in Mammalian Cells. *Analytical chemistry* 2017;89:9124-30.
152. Oyeyipo IP, van der Linde M, du Plessis SS. Environmental Exposure of Sperm Sex-Chromosomes: A Gender Selection Technique. *Toxicological research* 2017;33:315-23.
153. Larsson P, Andersson E, Johansson U, Ollinger K, Rosdahl I. Ultraviolet A and B affect human melanocytes and keratinocytes differently. A study of oxidative alterations and apoptosis. *Experimental dermatology* 2005;14:117-23.
154. Waster P, Orfanidis K, Eriksson I, Rosdahl I, Seifert O, Ollinger K. UV radiation promotes melanoma dissemination mediated by the sequential reaction axis of cathepsins-TGF-beta1-FAP-alpha. *British journal of cancer* 2017;117:535-44.
155. Lee JJ, Kim KB, Heo J, et al. Protective effect of *Arthrospira platensis* extracts against ultraviolet B-induced cellular senescence through inhibition of DNA damage and matrix metalloproteinase-1 expression in human dermal fibroblasts. *Journal of photochemistry and photobiology B, Biology* 2017;173:196-203.
156. Shalhout S, Haddad D, Sosin A, et al. Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development. *Molecular and cellular biology* 2014;34:4019-32.
157. Cabelof DC, Nakamura J, Heydari AR. A sensitive biochemical assay for the detection of uracil. *Environmental and molecular mutagenesis* 2006;47:31-7.
158. Lari SU, Chen CY, Vertessy BG, Morre J, Bennett SE. Quantitative determination of uracil residues in *Escherichia coli* DNA: Contribution of ung, dug, and dut genes to uracil avoidance. *DNA repair* 2006;5:1407-20.